The PHILOSOPHICAL MAGAZINE

FIRST PUBLISHED IN 1798

VOL. 41 SEVENTH SERIES No. 318 JULY, 1950

A Journal of Theoretical Experimental and Applied Physics

PRICE 10s.
Annual Subscription £5 2s. 6d. payable in advance.

PRINTED AND PUBLISHED BY TAYLOR & FRANCIS LTD., RED LION COURT, FLEET ST., LONDON, E.C.

CONTENTS OF No. 318.

*** All communications for the Philosophical Magazine should be addressed, post-paid, to the Editors, c/o Messrs. Taylor and Francis, Ltd., Red Lion Court, Fleet Street, London, England.

LVI. *The Thermal Conductivity of Potassium Chrome Alum at Temperatures below One Degree Absolute.*

By C. G. B. Garrett,
Royal Society Mond Laboratory, Cambridge*.

[Received April 12, 1950.]

Abstract.

An account is given of measurements of the thermal conductivity of potassium chrome alum made between 0·14° and 0·30° K., from which it is concluded that the mean free path of the phonons in the crystal is about $\frac{1}{2}$ mm. In the method employed, a long single crystal of the salt, isolated from the outside world, is demagnetized to a uniform low temperature: a temperature gradient along the crystal is then set up, and the time constant of the subsequent heat flow measured. The paper concludes with a consideration of the lowest temperature to which a sample of material could in a reasonable time be cooled by direct thermal contact with a cold mass of paramagnetic salt.

§ 1. Introduction.

Temperatures below 1° K. may be obtained by the process of magnetic cooling, in which a paramagnetic salt is used as the refrigerant. In order to extend the study of these very low temperatures to other materials it is desirable to have information about the time required for cooling by direct contact with the cooled salt, and one of the controlling factors in this process will be the thermal conductivity of the paramagnetic crystals themselves. In the following paper an account is given of measurements of the thermal conductivity of a typical paramagnetic salt, potassium chrome alum, in the range of temperature between 0·14° and 0·30° K. Such measurements are also of interest in connection with the theory of heat conduction in insulating materials.

Precise measurements of the thermal conductivity of potassium chrome alum have been made in the liquid helium temperature range by Bijl (1949), but below 1° only the order of magnitude of the conductivity is known (Kürti, Rollin and Simon 1936). The latter determination was made by demagnetizing a long single crystal of the salt from an inhomogeneous field, so that the ends were cooled to somewhat different

* Communicated by Sir Lawrence Bragg, F.R.S.

temperatures: the subsequent approach to temperature equality was watched by means of local magnetic susceptibility measurements made with mutual inductances placed around the two ends, and the conductivity estimated from semi-qualitative considerations. This method has the intrinsic advantage that it does not require separate thermometers or a heating coil attached to the crystal, thereby avoiding trouble from contact thermal resistances and heat inflow along metallic leads. In the present investigation a similar method has been used, in which however a device is employed for setting up a temperature gradient whenever required, without the necessity for a fresh demagnetization. In addition a more exact analysis of the heat flow has been made, and care has been taken to keep temperature differences as small as possible, with a view to investigating the temperature variation of the conductivity over the observable range.

§ 2. Principle of the Method.

The salt, in the form of a long single crystal, is demagnetized from a uniform magnetic field, so that its final temperature is uniform. A non-uniform magnetic field, obtained from a suitable system of coils, is applied to the crystal in such a way that the field varies linearly along the length of the crystal with a gradient of the order of 100 gauss/cm. A temperature gradient is set up and heat flows from the warmer to the cooler end of the crystal. After allowing time sufficient for the temperature to become once more approximately uniform, the non-uniform field is removed, and a temperature gradient is now set up in the opposite sense to the first. The thermodynamics of this process will be discussed in § 3. The temperatures near the two ends of the crystal may be measured by systems of mutual inductance coils suitably placed, and the rate of approach of those temperatures determined*.

From the rate of approach of temperature it is possible to calculate the diffusivity of the material, $K/\rho C$, where K is the thermal conductivity, ρ the density and C the specific heat per unit mass. The value of C appropriate to the mean temperature of the crystal may be taken from previous measurements, and hence the thermal conductivity may be derived. Since, in the temperature range above 0·2°, $C \sim A/T^2$, and since the conductivity itself is proportional approximately to T^3, the diffusivity, and hence the equilibrium time of the flow process, are extremely rapidly varying functions of temperature. It is therefore possible, for a given size of crystal sample, to make measurements only over a limited temperature range.

---

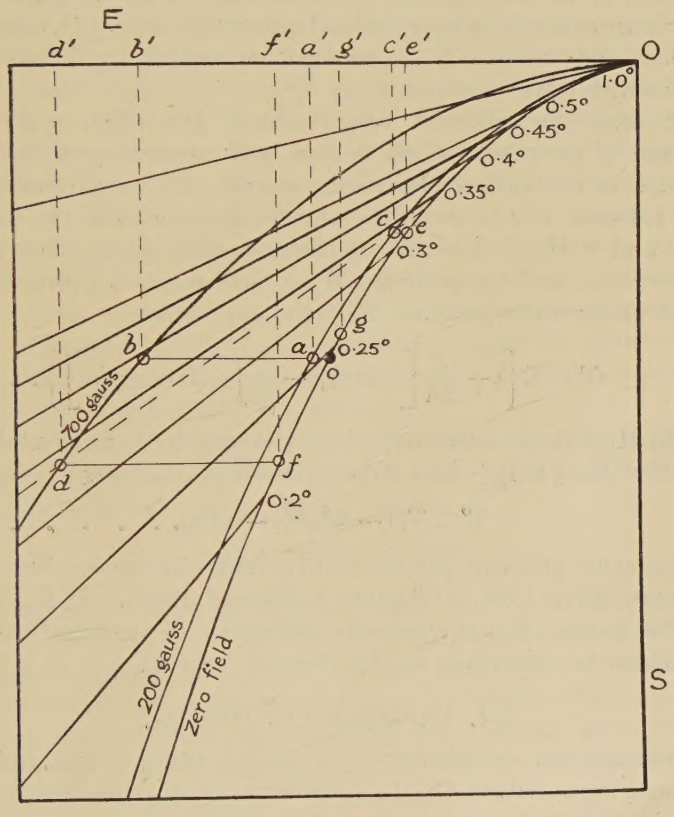
* In principle it would have been possible to do this during the course of the first heat flow, after applying the non-uniform field: in practice, however, fluctuations in such a field, to which the measuring system is very far from astatic, render accurate measurement impossible at this stage.

§ 3. Thermodynamics of the Temperature Gradient Initiation.

The principle of the device for establishing the temperature gradient may best be understood in terms of the Mollier diagram for the salt (fig. 1), in which the entropy S is plotted against the enthalpy E, defined so that $dE=T\,dS-M\,dH$. For the purpose of constructing such a diagram it may be assumed that the susceptibility follows a Curie law $\chi=\lambda/T$ and that the specific heat is of the form $C=A/T^2$, as is certainly a reasonable approximation throughout the temperature range considered. It may then be seen that "lines of constant field" are parabolæ, each one given by allotting to H a suitable value in the equation

$$S\simeq\frac{E^2}{A}\left(1-\frac{\lambda H^2}{A}\right). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Fig. 1.



"Lines of constant temperature", on the other hand, are straight, being given by the equation

$$E=2TS \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

in which the appropriate values are given to T. These lines must necessarily stop at the parabola corresponding to H=0.

Initially the crystal is assumed to be in a condition represented in fig. 1 by the point o. The initial magnetization in the non-uniform field is carried out isentropically, the two ends of the crystal reaching point such as a and b, in different fields and at different temperatures. In the subsequent heat flow the total enthalpy remains unchanged, so that the ends will reach points c and d on the same isotherm, determined by the condition that $a'c'=b'd'$. (These considerations should strictly be applied to all pairs of corresponding points in the two halves of the crystal.) On removing the non-uniform field isentropically the ends move to points e and f, representing temperatures respectively higher and lower than that of the initial state o. Having achieved such a gradient in the presence of zero magnetic field, the subsequent rearrangement of temperature may be studied experimentally and the diffusivity deduced.

After the second heat flow the final state of the system is represented by the point g, fixed by the condition that $e'g'=f'g'$. It may be verified that the temperature at g is somewhat higher than that at O, corresponding to a second-order increase in the entropy consequent upon the irreversible processes which have occurred.

To calculate the order of magnitude of the effect and the exact distribution of temperature set up, we may proceed analytically, using the arguments outlined qualitatively above. If x represents the axial distance between any point in the crystal and one end, the total length being $2l$, and if $\mathrm{H}(x)$ represent the field at that point, then it may be shown that the final temperature T' at that point is given in terms of the initial uniform temperature T by the expression

$$\mathrm{T}'=\mathrm{T}\left\{1-\frac{\lambda}{2\mathrm{A}}\left[\{\mathrm{H}(x)\}^2-\frac{1}{2l}\int_0^{2l}\{\mathrm{H}(x)\}^2\,dx\right]\right\}. \quad . \quad . \quad . \quad (3)$$

If the field $\mathrm{H}(x)$ varies uniformly along the length of the crystal following the equation $\mathrm{H}(x)=\mathrm{H}_0+hx/l$, it being assumed that $h\ll\mathrm{H}_0$, then

$$\mathrm{T}'\simeq\mathrm{T}[1-\lambda\mathrm{H}_0h(x-l)/l\mathrm{A}]. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

The temperature gradient produced is therefore uniform. For potassium chrome alum $\lambda/\mathrm{A}=1{\cdot}36\,.\,10^{-6}$ gauss^{-2} (Garrett 1948). If $\mathrm{H}_0=300$ gauss and $h=100$ gauss, then the extreme difference of temperature produced at 0·2° should be, according to this formula, 0·008°.

§ 4. Experimental Details.

The specimen was a single crystal of chrome alum, 65 mm. in length and of octagonal cross-section, the distance between opposite faces being about 15 mm. The crystal was cut in such a way that the axis of the sample coincided with one of the cubic axes of the crystal. It was suspended by fine nylon threads from a framework of german silver, which itself fits inside the conventional german silver salt-tube of a magnetic cooling cryostat. The measuring coils were wound on paper formers over the outside of the salt-tube, arranged so as to provide an estimate of the

averaged susceptibility over the last 15 mm. at each end of the crystal. Measurements of the two mutual inductances could be made alternately by comparison with a standard inductometer using a 40 cycle A.C. bridge network, detection of the balance point being made with the help of an amplifier and vibration galvanometer. In this way a temperature difference on the Curie scale of 0·01° at 0·2° could be determined with an accuracy of about 5 per cent.

The initial cooling was attained with a water-cooled armoured solenoid, which gives a maximum field of 10 Kg. After demagnetization, the non-uniform magnetic field required for setting up the temperature gradient was produced by (i) a solenoid of internal diameter 120 mm. and field-factor 36 gauss/amp. ; and (ii) a set of coils fitting inside the solenoid and around the liquid nitrogen dewar, designed to produce a linearly varying field of 10 gauss/amp./cm. A current of 8 amp. passed through these two coils in series for an appropriate interval of time was found to set up a temperature gradient of a suitable magnitude.

After switching off this non-uniform field, measurements of mutual inductance were made for the two coil systems alternately, the time of each determination being noted. Both sets of measuring coils were calibrated at temperatures in the helium range, so that mutual inductance readings could be used in each case to calculate T*. Unfortunately the corrections to T* for the shape of the crystal are unknown, and this uncertainty, although still moderately small at the lowest temperature investigated, will be a serious drawback in any attempt to extend the method to lower temperatures. As the field due to the primary of each of the mutual inductances is effective over roughly a spherical volume 15 mm. in diameter near the end of the crystal, it has been supposed that the measurements will give T^*_{sphere} directly.

It may be shown (see Appendix) that, to a good approximation, the difference in temperature between the ends of the crystal falls off exponentially with time according to the law

$$\Delta T \propto \exp(-\kappa\pi^2 t/4l^2), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

in which, as before, $2l$ is the overall length and κ the diffusivity of the crystal. In practice there will be, superimposed on these changes of temperature, a steady upward drift due to " heat leak ", which probably represents a surface recondensation of gaseous helium. Smoothed curves representing the terminal temperatures are drawn ; from this graph another curve is plotted, showing the difference in temperature plotted logarithmically against the time. The linearity of this graph checks the validity of the exponential law (equation (5)) and the diffusivity of the crystal at a mean temperature may be estimated from the gradient of the fitted straight line. The relation between T* and the true temperature T is known from Bleaney's measurements (Cooke 1949), as is also the specific heat at each temperature. The thermal conductivity itself may then be derived from the experimental values of the diffusivity.

§ 5. Results.

Measurements were made during three separate helium runs in October 1948. The results are given in the accompanying table.

October	T*	T	$4\kappa^2/l\pi^2$ seconds	C/R	K erg units	$\frac{10^{-5}K}{T^3}$
8	0·163	0·145	574	0·693	1·58	5·2
6	0·167	0·149	478	0·682	1·86	5·6
20	0·182	0·166	372	0·635	2·22	4·9
20	0·232	0·218	112	0·475	5·52	5·3
20	0·247	0·235	74	0·427	7·56	5·8
20	0·252	0·240	74	0·405	7·15	5·2
20	0·260	0·248	67	0·395	7·68	5·1
20	0·276	0·266	42	0·358	11·2	5·9
8	0·285	0·274	46	0·340	9·6	4·7

It will be seen that, as expected, the thermal equilibrium time does in fact depend very sharply on temperature. This, incidentally, explains why it has not so far been possible to extend the measurements to temperatures below 0·14° (where the state of transport of heat is no longer large in comparison with the heat leak) or above 0·30° (where an automatic recording device would have been necessary). The results are plotted against temperature in fig. 2, which shows that the conductivity is proportional to T^3 in this temperature range. It may be seen from the table that the average value of K/T^3 is

$$5{\cdot}3 \,.\, 10^5 \text{ ergs sec.}^{-1} \text{ cm.}^{-1} \text{ degree}^{-1} \quad . \quad . \quad . \quad . \quad . \quad (6)$$

with a mean deviation of $0{\cdot}12 \,.\, 10^5$. Owing to the possibility of systematic errors the actual uncertainty is rather greater than this.

Fig. 2 also shows the results of measurements made by Bijl (1949) in the liquid helium range, where the conductivity was found to follow a $T^{2\cdot3}$ law. Bijl found that the absolute magnitude of the conductivity appeared to depend on the rapidity with which the crystal was cooled from liquid nitrogen to helium temperatures. No attempt has been made to alter this rate of cooling during the present measurements, but it seems that our conditions correspond more closely with Bijl's " slow cooling ", and it will be seen from fig. 2 that a comparison of the two sets of measurements tends to confirm this. Kürti, Rollin and Simon (1936) estimated the thermal conductivity at 0·18° to be of the order of 10^{-4} watt units, and this also is in agreement with our results.

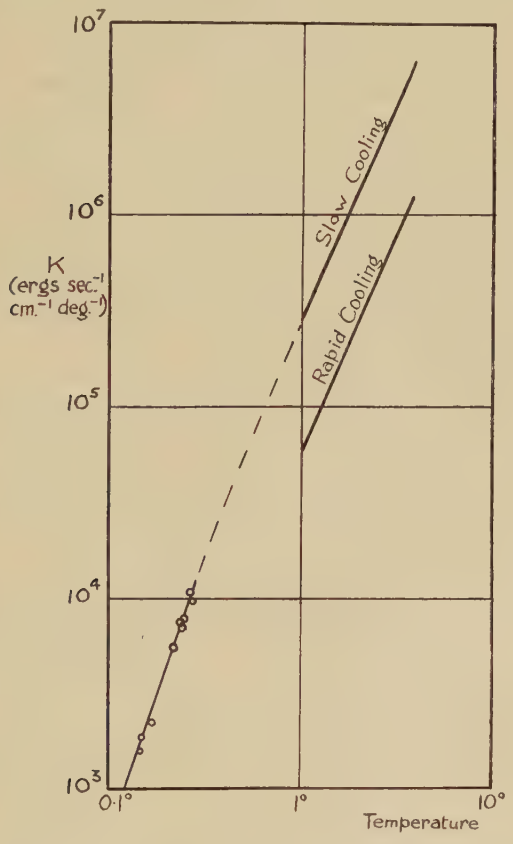
§ 6. Discussion.

In an insulating crystal the principal mechanism of thermal conduction is that associated with Debye waves or phonons. The phonon conductivity is given in order of magnitude by the expression

$$K_f \sim l_f s C_f, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

C_f being the lattice specific heat per unit volume, s the velocity of sound in the crystal and l_f the phonon mean free path. In this temperature range l_f should be a length specifying the separation between lattice boundaries, and so should be independent of temperature; s is also independent of temperature, and, since the lattice specific heat C_f should be proportional to T^3, it would be expected that the phonon conductivity would also be proportional to T^3.

Fig. 2.



In § 5 it was shown that a T^3 conductivity law is indeed followed for chrome alum at temperatures of the order of 0·2° K., although in the liquid helium range $K \propto T^{2\cdot3}$. Taking $C_f \sim 10^2 T^3$ ergs/° (Casimir 1940) and $s \sim 10^5$ cm./sec. we find that

$$l_f \sim 0{\cdot}05 \text{ cm.} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

a length considerably smaller than the macroscopic dimensions of the crystal (6·5 cm. × 1·5 cm.). Probably this is due to a mosaic structure in the crystal, as suggested by de Haas and Biermasz (1938) to account

for the "saturation" effect observed by them in experiments on the size dependence of the thermal conductivity of potassium chloride rods. Such a domain structure may also be associated with hysteresis effects in microwave experiments observed by Bleaney and Penrose (1948), and with the observation by Kraus and Nutting (1941) of a shattering of the crystal structure noticed on cooling a very similar material (ammonium chrome alum) through the temperature range 80–100° K.

Akheiser and Pomeranchuk (1944) have made the interesting suggestion that at the lowest temperatures the conductivity of paramagnetic crystals should again become large, due to the alternative mechanism for the transfer of thermal energy provided by the spin system ("exciton conductivity"). This is not in disagreement with the results presented in this paper, for calculation shows that at 0·14° K. the exciton conductivity should be some 5000 times smaller than the phonon conductivity. At somewhat lower temperatures the thermal equilibrium times become exceedingly long, but, if the predictions of Akheiser and Pomeranchuk are indeed correct, the diffusivity at 0·003° K. should be equal to that at 0·14° K., while at still lower temperatures it ought to increase as $1/T^2$. In view of the technical interest of such a possible increase in conductivity it is obviously of considerable importance that experiments be extended to this lower range of temperature.

From the information now available, it is possible to arrive at a tentative conclusion concerning the lowest temperature to which it is possible in a reasonable time to cool a sample of material by thermal contact with a cooled mass of chrome alum. It may be seen from equations (5) and (7) that the thermal equilibrium time within a crystal is proportional to $T^{-3}l^2l_f^{-1}$, where T is the temperature, l the macroscopic size of the sample and l_f the phonon mean path. In an actual experiment (see table of results) an equilibrium time of 500 seconds was found when $l \sim 65$ mm. and $T = 0{\cdot}15°$, l_f being subsequently deduced as 0·5 mm. From these figures it is possible to calculate what would be the equilibrium time at that temperature under the conditions of an actual contact cooling experiment; and hence to deduce how low the temperature could be before the equilibrium time would become intolerably long. In the experiments of Mendoza (1948) in which powdered salt was compressed around a copper insert under considerable pressure, an acetone-based lacquer being used as bonding agent, it was possible to reduce the distance between any grain of salt and the nearest copper surface to about 2 mm.; but on the other hand the phonon mean free path is probably given now by the size of the crystal grains, a distance very much smaller than the phonon mean free path of 0·5 mm. found for a large crystal. As a basis for comparison put $l = 1$ mm. and $l_f = 5 \cdot 10^{-3}$ mm. Using the above analysis it is found that at 0·05° K. the equilibrium time for such a sample should be of the order of 1000 seconds, which is about the longest time tolerable in a contact-cooling process. It may therefore be concluded that a temperature of 0·05° K. probably represents the limit

to which it would be practicable to cool a sample by direct thermal contact with a cooled salt. In this calculation no account has been taken of any surface thermal resistance between the salt and the metal. This assumption is based on the conclusion of Mendoza (1948) that below 0·2° K. the surface resistance is not the controlling factor.

§ 7. Acknowledgments.

I should like to thank Professor J. F. Allen and Dr. J. Ashmead for their help during the course of the work described in this paper. Thanks are also due to Mr. F. E. Smith of the University Chemical Laboratory, Cambridge, who supplied the crystal, and to the Department of Scientific and Industrial Research for a maintenance grant.

APPENDIX.

The exact solution of the thermal conduction problem, taking into account a diffusivity varying as T^5, would involve elaborate numerical computation. It is sufficient however to suppose the diffusivity constant at some mean value, and the problem is then soluble by standard methods (Carslaw and Jaeger 1947). Since radiation from the sides is of course negligible, the crystal may be treated as an infinite slab between two parallel planes, there being no heat flow across either boundary. If the initial temperature distribution be $f(x)$, then the temperature at some later time t at any point x is given by

$$T=\frac{1}{2l}\int_0^{2l} f(x')\,dx'+\frac{1}{l}\sum_{n=1}^{\infty} e^{-\frac{\kappa n^2\pi^2 t}{4l^2}}\cos\frac{n\pi x}{2l}\int_0^{2l} f(x')\cos\frac{n\pi x'}{2l}\,dx', \quad . \quad (9)$$

where κ represents the diffusivity of the material. Suppose a linear initial temperature distribution of the form $f(x)=T_0+s(x-l)$. Substituting in equation (9), all terms under the summation sign for which n is even vanish, and there remains the expression

$$T=T_0-\frac{8sl}{\pi^2}\cos\left(\frac{\pi x}{2l}\right)\exp\left(-\frac{\kappa\pi^2 t}{4l^2}\right)$$
$$-\frac{8sl}{9\pi^2}\cos\left(\frac{3\pi x}{2l}\right)\exp\left(-\frac{9\kappa\pi^2 t}{4l^2}\right)$$
$$-\ldots\ldots \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

It is possible, however, to neglect all except the first of the exponential terms, provided that we are not interested in the distribution at times very shortly after the initiation of the heat flow. Since the coefficient of t in the exponent in equation (10) is the same at all points along the

length of the crystal, it is immaterial whether temperature measurements are made over an infinitesimal region at the very end of the crystal, or whether instead an average is observed of the temperature near the end of the crystal. In either case

$$\Delta T \sim \exp(-\kappa\pi^2 t/4l^2), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

so that a graph of log T against t should be a straight line, from the gradient of which κ may be found.

References.

Akhieser, A., and Pomeranchuk, I., 1944, *Journal of Physics*, **8,** 216.
Bleaney, B., and Penrose, R. P., 1948, *Proc. Phys. Soc.*, **60,** 395.
Bijl, D., 1949, *Physica*, **14,** 684.
Carslaw and Jaeger, 1947, *The Conduction of Heat in Solids*, p. 85.
Casimir, H. B. G., 1940, *Magnetism and Very Low Temperatures*, p. 87.
Cooke, A. H., 1949, *Proc. Phys. Soc.*, **62,** 269.
de Haas, W. J., and Biermasz, Th., 1938, *Physica*, **5,** 47, 320, 619.
Garrett, C. G. B., 1948, *Cérémonies Langevin–Perrin, Paris* (in the press).
Kraus, D. L., and Nutting, G. C., 1941, *J. Chem. Soc.*, **9,** 133.
Kurti, N., Rollin, B.-V., and Simon, F., 1936, *Physica*, **3,** 266.
Mendoza, E. B., 1948, *Cérémonies Langevin–Perrin, Paris* (in the press).

ERRATA.

In the author's paper on *Quantal Aspects of Scientific Information*, which appeared in the March issue (p. 294), the relation between physical scales and proper scales is unfortunately obscured by an error in the example of §4(*c*).

The last two sentences on p. 294 should read : " We can now give a probability p of just $\frac{1}{2}$ to a statement in the form : x occupies $\sqrt{i}$ equal intervals. The figure of $\frac{1}{2}$ implies . . . . that its terminal is as likely to fall within the $\sqrt{i}$th equal as to fall outside it."

As pointed out in §4(*c*), the number of *conceptually*-separate occupance-relations, or metron-content, (§4(*b*)) is not the same thing as the number of *physically-separate* occupance-relations between a magnitude and the scale on which it is *observed* (§4(*a*)). Logically, i metrons enable us to add i elements to a formal representation of a measurement. The number of occupance-relations on the scale of *physical magnitude* graduated as in §4(*a*) may be quite different. In the example given, (§4(*c*)) $i=x^2/(\Delta x^2)$. But the physical scale-unit of x as defined in §4(*a*) is Δx, and if the probable error is constant, only $x/\Delta x$ equal intervals can be observationally separated on the scale of x, though i conceptually significant intervals can be defined. The probability is 1/2, not that the ith *conceptual* interval be occupied, as stated, but that the $\sqrt{i}$th *physical* scale-unit be occupied. The number of occupance-relations on the physical scale is, infact, of greater practical importance, though it lacks of course some of the invariant features of metron-content.

D. M. Mackay.

LVII. *The Crystallography and Thermodynamics of Order–Disorder Transitions in Certain Mixed Ketones.*

By Vera Daniel,
British Electrical Research Association*.

[Received March 20, 1950.]

[Plates XX. & XXI.]

Summary.

Previous work by Oldham and Ubbelohde (1939, 1940) on ketones of equal chainlengths and different positions of the ketogroup has shown that some of these compounds have practically identical crystal structures and physical properties; these will be called equivalent ketones. In mixtures of equivalent ketones all constituent ketones have identical carbon chains, and each ketone contains one ketogroup. Thus, mixtures of equivalent ketones differ from each other only in so far as the configuration of the ketogroups affects the bonding of their structure. In the solid state, the system formed by mixtures of different proportions of equivalent ketones is analogous to a crystalline compound, the bonding of which may be altered at will, within certain limits.

In the present paper the thermodynamic stability of different configurations of ketogroups in mixtures of equal amounts of equivalent ketones is treated theoretically. It is assumed the ketogroups contribute to the bonding only by virtue of their dipole moment.

Different mixtures of six equivalent ketones of chainlength eighteen were examined experimentally. They always form one of two structures, one orthorhombic (or monoclinic) and the other hexagonal, the hexagonal one being relatively disordered, as far as the packing of the chains is concerned. Two kinds of order–disorder transitions occur, one of which corresponds to a segregation of the individual ketones into separate phases, while the other one is analogous to the well known transitions in paraffins. The conditions of stability of the different configurations of ketogroups agree with the theory. In particular, the binding energy of a plane array of ketogroups can be derived from transition temperatures, and agrees with the electrostatic interaction energy derived for this array by Fröhlich (1946).

Introduction.

The relationship between the physical properties of substances and their chemical and crystallographic structure is difficult to investigate because their structure cannot be modified at will. A physical property

* Communicated by Sir Lawrence Bragg.

such as, say, dielectric constant or thermal conductivity, depends on a number of factors. The understanding and control of such a property would be easier if these factors could be altered separately and in small steps. However, this is usually not possible because a change from one chemical compound to another generally alters a number of structural features simultaneously, and in rather large steps.

Systems which permit the alteration of structural features in small steps can arise in two ways. Firstly, there exist series of compounds such that structural features within the series vary in a regular way and in more or less small stages. Such are the series of homologues in organic chemistry, for instance the paraffins, or the groups of analogous inorganic compounds such as the alkali halides. Here the modification of the structure is effected by chemical means and may be restricted to a more or less distinct structural feature, such as the length of a paraffin chain. Secondly, systems with variable structure may be produced by mixing two or more chemical compounds. Here changes can be made indefinitely small, but they are generally not restricted to any one particular structural feature.

In the present investigation both the above methods are applied simultaneously, in that the materials investigated are mixtures of closely related compounds, where both the original compounds and their mixtures have very similar crystal structures. These mixtures form a system which can in some ways be varied in small and quantitatively controlled steps. It may exist in forms of different degrees of order, and the conditions of stability of these forms can be derived thermodynamically, on the basis of certain simple assumptions.

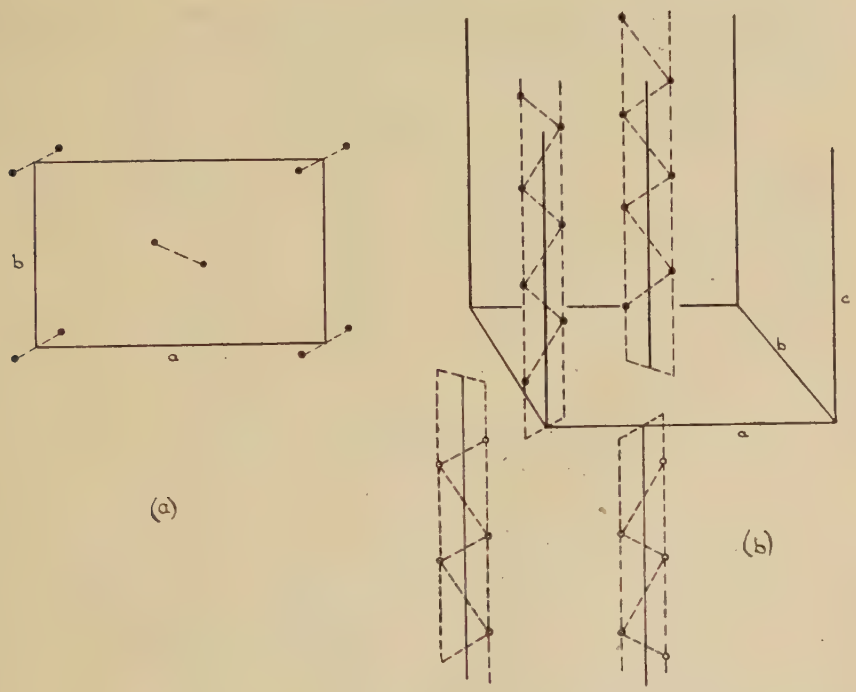
The Structure and Properties of Long-Chain Ketones.

The compounds used for the present work are long-chain ketones of the general formula $C_pH_{2p+1}CO\ C_qH_{2q+1}$, with p and q greater than 3. They differ from normal paraffins only in that two hydrogens are replaced by an oxygen. The C=0 group has a large dipole moment, and a ketone of this type may be considered as a paraffin to which a dipole has been rigidly attached, approximately at right angles to the chain.

The structure of normal long-chain paraffins has been fairly fully investigated in the past. Müller (1928, 1930) determined the crystal structure of the paraffin $C_{29}H_{60}$ (see fig. 1). This structure is typical for the long-chain paraffins and many of their derivatives. Its main features are that each molecule forms a comparatively rigid flat zig-zag chain and that the chains are parallel to each other in layers. A projection of a layer of chains on the basal plane shows a face-centred arrangement (fig. 1(*a*)). There exist different modifications of this structure. The chains may be normal or inclined to the basal plane, and the fit of the chain-ends may vary.

Long-chain paraffins show a variety of polymorphic transitions (Piper, Chibnall and Collaborators 1931). Many of these transitions are connected with changes of the angle between the chain axes and the basal plane of the structure, while in other cases the chains are normal to the basal plane throughout the transition. The latter type of transition was investigated by Müller (1932) with the help of X-rays. Müller found that the transition consists in a very anisotropic expansion of the basal plane ab (fig. 1 (*a*)) such that the chains in it approach a hexagonal arrangement. That is the cross section of the chains becomes statistically circular rather than flat. This change of structure is gradual in some paraffins and discontinuous

Fig. 1.



The structure of the paraffin $C_{29}H_{60}$. (Simplified from Müller 1928.)

in others. It represents a typical order–disorder transition, as is also confirmed by the calorimetric data of Garner, Bibber and King (1931) and Ubbelohde (1938). Dielectric measurements by Jackson (1935a, b) Müller (1936, 1939) Pelmore (1939) and Sillars (1939) on paraffins with dipolar groups indicate that in some paraffin lattices the chains have considerable freedom to rotate and/or twist.

As regards ketones as distinct from paraffins, a number of ketones $C_pH_{2p+1}CO\,C_qH_{2q+1}$ can be produced for any one long-chain paraffin of a given chainlength $p+q+1$. Oldham and Ubbelohde (1939) measured the melting points of all ketones of chainlengths seventeen and eighteen, and

in addition measured the latent heat of melting and entropy of melting of all ketones of chainlengths seventeen. Table I. gives their results in a slightly simplified form. It shows that for all ketones of the same chainlengths, with p and q greater than three, all the measured quantities are very similar. The only exception is $C_8H_{17}CO\,C_8H_{17}$ which differs from the other ketones in that its ketogroup is symmetrically placed in the centre of the chain. All the unsymmetrical ketones of equal chainlengths and with ketogroups away from the chain-ends have closely similar latent heats and entropies of melting and melting points.

TABLE I.

Ketones of total chainlength $p+q+1=17$.

p	T_M °C.	ΔH k cal./mole	ΔS cal./mole/°
1	47·8	14·1	44·0
2	47·6	13·9	43·2
3	41·6	12·7	40·4
4	41·5	12·7	40·3
5	40·9	12·7	40·5
6	40·6	12·9	40·9
7	41·8	13·1	41·7
8	50·0	13·7	42·3

Ketones of total chainlength $p+q+1=18$

p	T_M °C.
1	51·5
2	51·0
3	45·8
4	45·2
5	44·8
6	45·0
7	44·6
8	45·8

T_M ....melting point,
ΔH ....latent heat of melting,
ΔS ....entropy of melting.

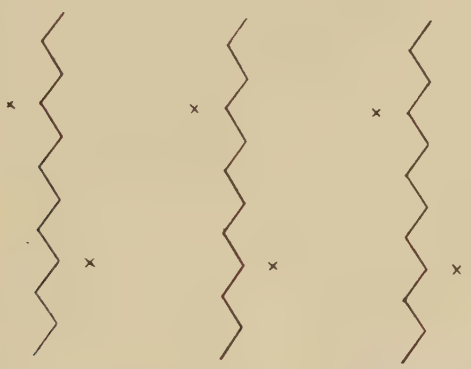
Oldham and Ubbelohde (1939) examined the ketones of chainlengths seventeen by X-rays and microscopically without finding significant differences between them. A similar examination was carried out by the author for ketones of chainlengths eighteen*.

* The author is greatly indebted to Dr. Oldham of the Royal Institution for the loan of small samples of five ketones of chainlengths eighteen with $p=3$, 4, 5, 7 and 8. A sixth ketone, with $p=6$, was synthesized according to Dr. Oldham's advice.

X-ray powder photographs were taken of six ketones of that chainlength and $C_6H_{13}COC_{11}H_{23}$ was investigated in somewhat greater detail, using single crystals; but no complete structure determination was made. All the ketones have unit cells of orthorhombic shape, and crystallize with their chain axes normal to the basal plane. (This does not, of course, exclude a monoclinic space-group, and that of $C_6H_{13}COC_{11}H_{23}$ is probably C_{2h}^6 or C_s^4.) No significant difference was found between the unit cell dimensions of the six ketones and inspection of the powder photographs indicates that in all the ketones the chains are arranged very similarly. The X-ray evidence will later be discussed in more detail.

The structure of an unsymmetrical ketone of the kind discussed here has some interesting features which have been considered by Fröhlich (1946)*. As fig. 2 shows, the ketogroups belonging to a layer of chains may be situated at one of two levels. Within either of these two levels

Fig. 2.



Schematic representation of a layer in a crystalline ketone. The positions which may be occupied by oxygen atoms are marked by crosses.

* Fröhlich's paper is mainly concerned with a theory of order–disorder transitions in pure ketones. Müller's (1936, 1938) dielectric measurements for the ketones $C_8H_{17}COC_8H_{17}$ and $C_{12}H_{25}COC_{12}H_{25}$ appeared to indicate the inception of such a transition near the melting point: they show the dielectric constant to change gradually, over a temperature range, from the value for the solid to that for the liquid. This gives the impression that a transition would take place at a temperature above the melting point if the material did not melt before this temperature is reached. However, dielectric measurements on very carefully purified samples of $C_8H_{17}COC_8H_{17}$, $C_6H_{13}COC_{11}H_{23}$ and $C_5H_{11}COC_{11}H_{23}$ indicate that in pure solid ketones the dielectric constant is almost independant of temperature, and that there is no gradual transition from the value for the solid to that for the liquid. The change from solid to liquid in a pure ketone is discontinuous at the melting point. and the gradualness of the transition found by Müller appears to be due to small traces of impurities. Thus there is no evidence that pure ketones show order–disorder transitions. The dielectric measurements mentioned will be described elsewhere.

the ketogroups form a planár arrangement, such that the dipoles represented by the ketogroup are roughly parallel to each other. Fröhlich shows that this arrangement is very favourable from the point of view of electrostatic interaction. If a level is fully populated, so that each chain contributes a dipole to it, the electrostatic interaction of the dipoles contributes quite a large term to the binding energy of the structure.

The arrangement of the paraffin chains in the structure leaves it undecided whether ketogroups shall be concentrated in "dipolar planes" or distributed at random between the two possible levels. However, the two possibilities imply different energy conditions. The random arrangement has a higher entropy. On the other hand, the electrostatic interaction is stronger if the ketogroups are concentrated at one level, because the electrostatic interaction energy between dipoles varies inversely with the third power of their distance. If the dipoles are distributed between two levels the numerical value of the total dipolar interaction energy falls to about a quarter of that for a fully populated dipolar plane. If the dipoles are concentrated in dipolar planes the structure will have a lower entropy and a lower energy than if they are distributed at random between two levels, for the dipolar interaction reduces the total energy H of the structure by restricting some of the degrees of freedom. Thus the free energy $G=H-TS$ of the form with dipolar planes has a lower H and a lower S than the random form. Whether or not it has a lower free energy depends on the numerical value of the energy of the dipolar plane. This energy is, as it were, balanced against the entropy term. If it is sufficiently large, the form with dipolar planes will be stable. The thermodynamic conditions of stability in question will be discussed in more detail later.

It is a matter for experiment to decide which of the possible forms of an unsymmetrical ketone is in fact formed. From the point of view of X-ray crystallography the two forms, with and without dipolar planes, differ very little from each other. A preliminary investigation of single crystals of the ketone $C_6H_{13}COC_{11}H_{23}$ showed that it would need a detailed determination of the crystal structure to decide which of the two possible forms is present. However, the two arrangements differ radically in their ability to form solid solutions with other, similar, ketones, as will be seen from the following.

Mixtures of Equivalent Ketones.

As Oldham's and Ubbelohde's work (1939) shows, unsymmetrical ketones of equal chainlengths but different position of the ketogroup have practically identical crystal structure and thermal properties unless their ketogroups are near the chain-ends. These ketones, of the general formula $C_pH_{2p+1}COC_qH_{2q+1}$ with $p \geqslant 3q \geqslant 3$ and $p \neq q$, will be called equivalent ketones. Because of the similarity of the "equivalent" ketones one would expect them to form mixed crystals which were very

nearly ideal solid solutions. However, this cannot be the case if the ketones form dipolar planes with a high energy of electrostatic interaction, because in a solid solution these planes must be broken up. Consequently, if the energy of a dipolar plane is large, two equivalent ketones may be mutually immiscible, in spite of their similar structure. On the other hand, if no dipolar planes are formed in the pure equivalent ketones, they should form solid solutions with each other.

Oldham and Ubbelohde (1940) in a study of structures with flaws, measured the melting points of a number of binary mixtures of equivalent ketones. Their results show that in thermodynamic equilibrium these ketones do not mix in the solid to any appreciable extent. This indicates that the pure equivalent ketones crystallize with their ketogroups segregated in dipolar planes.

The two-phase character of binary mixtures of equivalent ketones shows that the energy of dipolar planes must be rather large for the following reason. In pure ketones there exist two levels among which the ketogroups might be distributed at random, while the configuration of the paraffin chains is retained. (See fig. 2.) In a mixture of two equivalent ketones with different positions of the ketogroup there are four such levels, and the formation of dipolar planes means a greater loss of randomness than in a pure ketone. This means that in a mixture of two equivalent ketones the energy of the dipolar planes has to outweigh a larger entropy term if it is to enforce the segregation of dipoles in dipolar planes. The fact that the segregation does occur in binary mixtures implies that the energy of the dipolar planes is sufficient even for that. However, what will happen on mixing three, four, or more equivalent ketones? In these cases there will be still more levels available among which the ketogroups can be distributed at random, and there must be a number of ketones for which the energy of the dipolar plane will be insufficient to enforce the segregation of the dipoles and a solid solution will be formed.

The Thermodynamics of Mixtures of Equivalent Ketones.

The qualitative reasoning in the preceding paragraph may be made quantitative, subject to some simple assumptions. It may however be useful first to restate the problem in somewhat different terms.

There exist many substances which crystallize at low temperatures in an ordered structure which changes to a disordered structure as the temperature rises. In all these cases an energy term—making for order—militates against an entropy term which makes for disorder. The entropy term contains the temperature as a factor while the energy term does not and so the entropy term must outweigh the energy term at some sufficiently high temperature.

The case of two substances which in the solid state tend to segregate is analogous to an order–disorder transition. Here again an energy

term makes for segregation while an entropy term makes for solid solution. Two such substances are said to form a solid solubility gap which closes at some temperature, so that above that temperature a solid solution is always formed.

Mixtures of equivalent ketones have the characteristics both of a single substance undergoing an order–disorder transition and of a number of substances forming a multiple solid-solubility gap. The fact that we have a number of equivalent substances makes it possible to increase the entropy term—which makes for disorder and solid solution—in discrete steps. The disordered form may be made stable by increasing the number of ketones instead of the temperature.

In the following we consider mixtures of equal amounts of n equivalent ketones and deduce their conditions of thermodynamic equilibrium. To this end we make the following assumptions :

1. The ketones considered have equal total energies and entropies for equal temperatures and volumes when in equilibrium.

2. Liquid mixtures of the ketones are ideal solutions.

The first assumption depends for its justification on the equality of the latent heats and entropies of melting of ketones of chainlength seventeen (Oldham and Ubbelohde 1939) and the similarity of the structure of the ketones. The second assumption implies that in liquid ketones the ketogroups are distributed at random.

For solid mixtures of n ketones, Gibbs' phase rule implies that the mixture can consist of not more than n phases in equilibrium (except at a so-called $n-2$ fold point). With regard to the nature of the possible phases we make the assumption :

3. Any ketone in the solid mixture may crystallize either with or without dipolar planes. That is, the ketogroups may either all be concentrated in one level or else distributed at random between all the $2n$ possible levels. This assumption implies that the solid mixture may be either n-phase or single-phase, since if the form with dipolar planes is stable for one ketone, it must be stable for all the ketones. This, because equal amounts of the n ketones are present and thus the specific volume is equal for all of them and because they are equivalent as defined in the first assumption. If the form with dipolar planes is stable the mixture will be n-phase, each phase consisting predominantly of one ketone. If the form without dipolar planes is stable the ketones will form a single-phase solid solution. We shall make one more simplifying assumption with regard to these two possibilities :

4. If the mixture is single phase, it is a completely random solid solution. If it is n-phase, the mass of each ketone is concentrated in its " own " phase, the remainder of its mass which is distributed among the other $n-1$ phases being negligible.

In principle, conditions are likely to exist in which each ketone contains a fair quantity of the other ketones in solid solution, while still forming a separate phase. However, such conditions are considered to be

restricted to narrow ranges of temperature so that they may be neglected. This is equivalent to assuming the transition from the form with dipolar planes to the form without dipolar planes to be a transition of the first order.

In a mixture of equal amounts of n equivalent ketones all the ketones behave alike and so it is sufficient to write down the thermodynamic functions for one of them.

For a pure ketone we put the Gibbs free energy per mole in the solid state

$$G_0=H_0-TS_0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

and in the liquid state

$$(G_0)=(H_0)-T(S_0) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

the notation used is that of Oldham and Ubbelohde (1940). G is defined for constant pressure so that H contains a volume term.

As the liquid mixture of n equivalent ketones is assumed to be an ideal solution, the partial total energy of each equivalent ketone is the same as for a pure liquid ketone, but the partial entropy is increased by a term equal to R log n because the volume available for each ketone is increased n fold (see for instance Fowler and Guggenheim, *Statistical Thermodynamics*, p. 161). Thus for the liquid mixture of n ketones

$$(G)=(H_0)-T((S_0)+R \log n). \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

The solid mixture may be either n-phase or single phase. If it is n-phase the assumptions made imply that all except a negligible part of each ketone is in the same state as in a pure ketone.

Thus $$G_1=H_0-TS_0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

If the solid mixture is single-phase, the partial entropy of each ketone will be increased by a mixture term, which is again R log n if the ketones form ideal mixed crystals (see Fowler and Guggenheim, *Statistical Thermodynamics*, p. 185). In addition there will be an entropy increase of R log 2, due to the fact that in a pure unsymmetrical ketone all dipoles are on one level, while in the mixture the dipoles of any one unsymmetrical ketone are on two levels. Apart from this configurational increase of entropy R log $2n$ there may be present a further increase of entropy because the removal of the dipolar planes is likely to make the whole structure looser. This dynamic part of the entropy is difficult to estimate. We shall treat it as an unkown constant s, independent of n. This is most likely an over-simplification, as the dynamic part of the excess entropy probably depends on the number of levels on which dipoles may occur, and thus on n, but it seems reasonable to neglect this dependence, in first approximation. The total energy H of the mixture will be increased by the energy of the dipolar plane h_D, and if there is a further loosening of structure beyond the removal of the dipolar planes,

it may be increased by a further term h_x. Both these terms are independent of n, in first approximation. This makes the free energy of the single-phase mixture

$$G_2=H_0+h_D+h_x-T(S_0+R \log 2n+s). \quad . \quad . \quad . \quad . \quad (5)$$

The assumptions implicit in this equation mean that the single phase mixture may have another structure than the pure ketone—hence h_x and s—but that the thermal characteristics of this structure are independent of n.

The equations which have been deduced permit calculation of the melting points of the n-phase and single phase mixtures. For the melting point T_1 of the n-phase mixture $G_1=(G)$, so that

$$(H_0)-H_0=T_1((S_0)-S_0)+RT_1 \log n$$

where $(H_0)-H_0=\Delta H$ is the latent heat of melting of a pure ketone and $(S_0)-S_0=\Delta S$ its entropy of melting. Using the equation $\Delta H=T_0\Delta S$ valid for the melting point T_0 of the pure ketone, the melting point of the n-phase mixture is given by

$$\Delta H\left(\frac{1}{T_1}-\frac{1}{T_0}\right)=R \log n, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

Equation (6) is simply the Hildebrand equation for the melting points of inmiscible substances, written for the special case of equal amounts of substances of equal ΔH and ΔS.

The melting points of the single phase mixtures follow from $(G)=G_2$ and should obey the equation

$$\Delta H\left(\frac{1}{T_2}-\frac{1}{T_0}\right)-\frac{h_D+h_x}{T_2}=R \log 2+s. \quad . \quad . \quad . \quad . \quad (7)$$

This implies that T_2, the melting point of the single phase mixtures, should be independent of n.

Comparison of equations (6) and (7) shows that with increasing n the melting point of the n-phase mixtures decreases while that of the single phase mixtures remains constant. This implies that even if the multiphase mixture is thermodynamically stable up to the melting point for low values of n, it must become unstable as compared with the single-phase form if n is made large enough. The condition in which the two forms are in equilibrium at a temperature T can be derived by putting $G_1=G_2$ which leads to the equation

$$h_D+h_x=T(R \log 2n+s). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

The lowest value for n at which the single phase form is stable in the solid is obtained if T is the melting point of the single phase form, that is, if $T=T_2$.

Melting Points of Mixed Ketones.

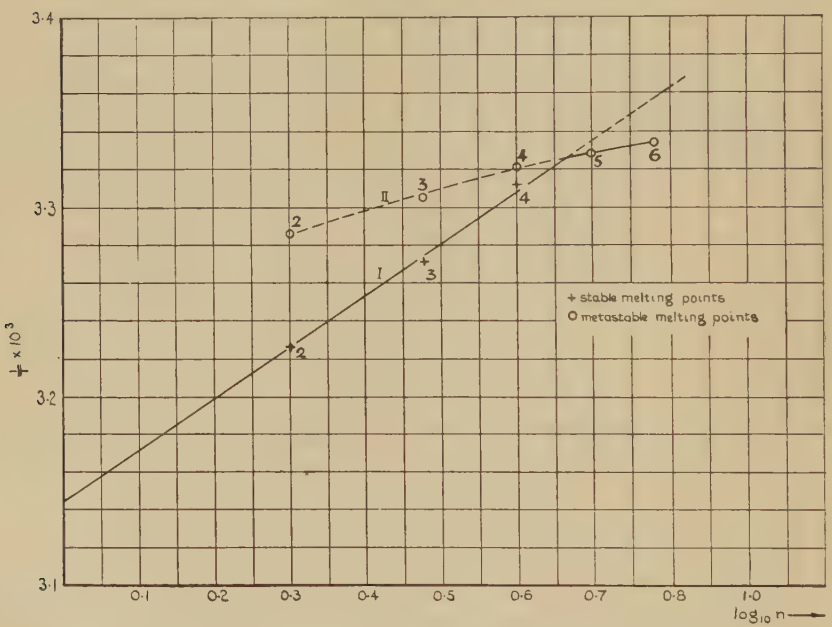
Equations (6) to (8) derived in the last paragraph may be compared with experiment by measuring the melting point of mixtures of equivalent ketones. By the courtesy of Dr. Oldham, the author had available small samples of all six equivalent ketones with chainlengths eighteen. Mixtures were prepared of equal amounts of these, and their melting points were measured in the usual way, the mixtures being held in glass capillaries. These measurements led to the observation that mixtures of 2, 3 and 4 ketones can exist in a metastable state. If the mixtures are allowed to solidify slowly, and kept at room temperature for some time, they reach a stable condition to which corresponds a certain melting point. If, however, they are quenched in cold water and then quickly immersed in warm water, they melt at a lower temperature, a metastable melting point. If the liquid mixture is kept for some time at a temperature between the metastable and the stable melting points, it ultimately solidifies again in the stable form. These observations show that, on quenching, the mixtures in question solidify in a mestastable form which changes in the solid state into the stable form if given time. The persistence of the metastable form varies markedly with the number n of ketones in the mixture. Mixtures of four equivalent ketones persist in the metastable state for several days, while in mixtures of two ketones the metastable melting point could only be observed if the quenching in water and ice was followed very quickly by immersion in warm water. No metastable melting points were observed in single ketones.

There exist in general several combinations of equal amounts of n ketones for any number of n. With six equivalent ketones of chainlength eighteen, the number of combinations is, respectively, 15, 20, 15, 6 and 1 for mixtures of 2, 3, 4, 5 and 6 ketones. It would have needed prohibitive amount of material to make all possible mixtures. Three combinations were made of four ketones, and it was found that the values measured for the stable and the metastable melting points did not depend noticeably on the individual ketones making up the four fold mixtures. Variations found were within a range of 1° C. For other mixtures only one or two combinations were used for each value of n.

The results of the measurements are shown in fig. 3. To facilitate comparison with the theoretical equations 6–8, the reciprocals of the melting points are plotted against the logarithm of n. The theory postulates that 1/T should vary linearly with log n if the mixtures are n-phase, while for single-phase mixtures 1/T should be independent of n. It may be seen that for $n=1$ to 4 the stable melting points lie on a straight line (I). The metastable melting points depend much less on n and lie on another curve (II), together with the stable melting points of the mixtures of five and six ketones. The experimental results can be interpreted in terms of the theory, if it is assumed that curve I corresponds to n-phase mixtures, while curve II corresponds to single-phase mixtures. This means that up to $n=4{\cdot}6$ where the two curves

intersect, the n-phase state is thermodynamically stable while the single phase state is stable for higher values of n (at the melting point). The existence of metastable single-phase mixtures can be easily understood because the formation of a multi-phase mixture from the melt involves extensive sorting out of the mixed ketones. This segregation is a process of diffusion and needs time. Thus, if the mixture is quenched it has to solidify in the random single-phase form. This interpretation explains why mixtures of four ketones persist in the metastable state for several days, while mixtures of two ketones become stable in a few seconds (the temperature being equal in both cases). The segregation of four ketones should need more time than the segregation of two.

Fig. 3.



Melting points of mixtures of n equivalent ketones.

According to equation (6) the slope of curve I should give ΔH, the latent heat of melting of a pure ketone. The value found from the experimental data is $\Delta H = 16{\cdot}6k$ cal/mole. Oldham and Ubbelohde (1940) give the latent heat of melting of one of the equivalent ketones used in the present work, namely $C_5H_{11}CO\,C_{12}H_{25}$, as $\Delta H = 14/6k$ cal/mole. The agreement is satisfactory, as equation (6) embodies some assumptions which are unlikely to be strictly true. The value of $R \log n$ for the entropy of mixing of the liquid mixture of ketones is an upper limit, and this term will be less if assumption 2 is not obeyed and the liquid mixture of ketones is not an ideal solution. If that is the case the value of ΔH deduced from equation (6) is spuriously large. Similarly, it is

presumably an exaggeration to assume that the ketones are entirely immiscible in the solid multiphase form. If they are somewhat soluble in each other, equation (6) will again give a spuriously high value of ΔH. As both the factors mentioned tend to make ΔH too large, the agreement found shows that the assumptions made are reasonably correct.

Equation (7) postulates that the melting points of the single-phase mixtures should be independent of n. Experiment shows that these melting points do not depend very strongly on n. However, they are not constant; mixtures of fewer ketones have higher melting points than mixtures of larger numbers of ketones. This is not surprising, for the constancy of h_x, h_D and s in the equations could hardly be expected to be more than a crude first approximation. Some dipolar interaction must occur even in random mixtures and this will be more marked the more densely populated the levels containing dipoles. Dipolar interaction should increase the melting points of single-phase mixtures of few ketones, where the levels are relatively densely populated, in comparison with mixtures of many ketones.

The Crystal Structures of Mixed Ketones.

The evidence of the melting points indicates the existence of two forms of the mixed ketones, a multi-phase form and a single-phase form. It remains now to investigate the structures of these forms, with a view to interpreting the energies and entropies involved in the order–disorder transition.

X-ray powder photographs were taken of a number of mixtures of the six equivalent ketones of chainlengths eighteen, both in the stable and the metastable state. As the most important Debye–Scherrer lines of the ketones occur at Bragg angles of 2 to 20°, an experimental arrangement had to be used which gave sharp lines for these angles. A Unicam single-crystal gonimeter was used, in conjunction with a slit collinator, the slits being about 0·2 mm. wide. The specimens were enclosed in cellophane tubes. For the measurement of spacings flat film was used, the specimen-film distance being 10 cm., and for a survey of the diffraction pattern a cylindrical filmholder of 3 cm. radius.

The results of the X-ray examinations were as follows: In all cases when according to the melting point data an n-phase mixture was expected, its X-ray diffraction pattern was consistent with a multi-phase mixture of the component ketones (see Pl. XX.(*b*)). In all cases when a single-phase mixture was expected, it had an hexagonal structure, different from that of the pure ketones (see Pl. XX.(*a*)). The structures of all single-phase mixtures were identical, within the limits of error. In the following these results will be discussed in greater detail.

The six equivalent ketones have closely similar X-ray patterns in that they have identical unit cells and the intensities of most Debye–Scherrer lines are very similar. The reason for this is that all the

ketones have identical paraffin chains which are apparently arranged very similarly in the different ketones. As the ketones contain only one oxygen atom for a paraffin chain of eighteen carbon atoms, the intensities of the Debye–Scherrer lines are mostly determined by the paraffin chains. Only if the structure amplitude of the paraffin chain for an X-ray reflection is small does the position of the ketogroup become important for the intensity of that reflection. This happens to be the case for the $00l$ reflections with l up to about 17, and as the $00l$ reflections up to $l=5$ are clearly distinguishable on powder photographs, it is easily possible to identify a ketone from a powder photograph. Otherwise, however, the X-ray diffraction patterns of the ketones are essentially due to their paraffin chains.

The multi-phase mixtures have orthorhombic unit cells, the dimensions of which are closely similar to those of the pure ketones (see Table II. *a*), The signficance of the small discrepancies found is not easy to assess, because, as has been pointed out by McArthur (1944), even with exceedingly pure ketones and paraffins unit cell dimensions are not reproducible within the geometric limits of accuracy. The powder patterns are too complex to be analysed in detail, but as far as visual examination can show, the diffraction patterns of the multi-phase mixtures are very similar to those of the pure ketones with regard to the intensities of the reflections other than the first five $00l$ reflections. This means that the arrangement of the paraffin chains in the multi-phase mixtures is very similar to that in the pure ketones. This agrees with the prediction that these mixtures should be multi-phase mixtures of the pure ketones.

As for the postulate of the theory that the individual ketones in the multi-phase mixtures should form separate aggregates with dipolar planes, the X-ray data are insufficient to give a conclusive answer. The only reflections which are different for the individual ketones, and could be identified on powder photographs, are the first five $00l$ reflections. The intensities of these reflections are only little dependent on whether the individual ketones form separate aggregates. However, even if more data were available, the investigation of this problem would meet the following difficulty.

In general, multi-phase aggregates of crystals may be expected to scatter X-rays incoherently, because the regularity of the repeat of the crystal structure will normally be interrupted at the boundaries between aggregates. However, this need not be the case with a multi-phase mixture of equivalent ketones, where the individual ketones have identically arranged paraffin chains. Aggregates of the different ketones, each containing dipolar planes, may be packed together so as to scatter coherently, as if the whole conglomerate were one single crystal. If this were the case, only a very detailed X-ray investigation could decide whether dipolar planes are formed or not. Such an investigation would be very interesting from the point of view of X-ray optics, but has so far not been undertaken.

TABLE II. *a*.

Cell dimensions of mixtures at room temperature

Mixture	*n*	*a*	*b*	*c*	Remarks
$C_3+C_4+C_5+C_6+C_7+C_8$	6	8·22±0·04	4·74±0·04	25·16±0·15	Stable, hexagonal
$C_3+C_4+C_5+C_6+C_7$	5	8·20±0·04	4·73±0·04	24·76±0·15	,, ,,
$C_4+C_5+C_6+C_7$	4	8·18±0·04	4·72±0·04	24·68±0·15	Metastable, hexagonal
$C_4+C_5+C_6+C_7$	4	7·60±0·04	4·92±0·04	24·94±0·15	Stable, orthorhombic
$C_3+C_5+C_6+C_7$	4	7·60±0·04	5·01±0·04	24·82±0·15	,, ,,
$C_3+C_5+C_6$	3	7·60±0·04	4·89±0·04	24·88±0·15	,, ,,
C_5+C_6	2	7·60±0·04	4·89±0·04	24·92±0·15	,, ,,
Average of six pure ketones	1	7·59±0·02	4·87±0·02	24·78±0·06	,, ,,

TABLE II. *b*.

Mixture	*n*	*a*	*b*	*c*	Remarks	T (°C.)
$C_3+C_4+C_5+C_6+C_7+C_8$	6	7·64±0·04	4·95±0·04	24·68±0·15	Stable, orthorhombic	4
		7·66±0·04	4·97±0·04	24·83±0·15	,, ,,	11
$C_3+C_4+C_5+C_6+C_7$	5	7·56±0·04	4·98±0·04	24·68±0·15	,, ,,	7·4
$C_4+C_5+C_6+C_7$	4	7·56±0·04	4·95±0·04	24·94±0·15	Metastable orthorhombic	7°
		7·54±0·04	5·00±0·04	24·82±0·15	Stable, orthorhombic	6°
$C_3+C_5+C_6+C_7$	4	7·56±0·04	4·80±0·04	25·16±0·15	,, ,,	7°

NOTE : *a*, *b* and *c* in Angstrom units. C_ρ indicates a ketone of formula $C_pH_{2p+1}COC_{17-p}H_{2(17-p)+1}$. Mixtures consist of equal amounts of the *n* ketones.

The above considerations show that the X-ray evidence on the multi-phase mixtures is not quite conclusive. However, the evidence of the unit cell dimensions and the arrangement of the paraffin chains confirms the multi-phase character of these mixtures, and the intensities observed are consistent with it. This appears to establish a strong case for the conclusion drawn from the melting point data, that the stable mixtures of up to four ketones consist of aggregates of the individual ketones, each forming dipolar planes.

As for the single-phase mixtures, X-ray photographs were obtained of stable mixtures of five and six ketones and of metastable mixtures of four ketones; metastable mixtures of three and two ketones were not photographed, because of their short life. All these mixtures gave X-ray patterns corresponding to a structure with an hexagonal arrangement of chains in the basal plane $a\,b$. The structure of the pure ketones is orthorhombic (or alternatively monoclinic with a monoclinic angle of very nearly 90°) and the chains in the basal plane $a\,b$ are arranged in a distorted hexagon (see fig. 1(a)). For the single-phase mixtures the distortion of this hexagon vanishes so that $a=\sqrt{3}(b)$.

The unit cell of the single-phase form may be referred to hexagonal indices, when the reflections 110 and 200 of the orthorhombic notation coalesce into one reflection, with indices 100 in the hexagonal notation. This reflection which is shown on Pl. XX.(a), and another one, corresponding to the coalescing of the orthorhombic 310 and 020 into the hexagonal $2\bar{1}0$ were the only $hk0$ reflections which could be distinguished on the powder photographs. Even the second line, the hexagonal $2\bar{1}0$, was so weak that monochromatic radiation had to be used for its detection. This indicates strong thermal vibrations at right angles to the chains. However, the reflections present establish a hexagonal packing of the chains in which each is surrounded by six nearest neighbours—the closest packing for cylindrical rods.

The $00l$ reflections of the single-phase mixtures correspond to a c axis equal to that of the pure ketones and multi-phase mixtures (see Table II.a). This implies that the paraffin chains are at right angles to the basal plane, as in the pure ketones. The intensities of the $00l$ reflections of any single phase mixture of four ketones are closely similar to those of the multi-phase mixture of the same four ketones. As has been pointed out above, the intensities of these reflections cannot be used to determine whether the ketogroups are distributed at random or arranged in dipolar planes. It may be noted in this context that the high orders of the $00l$ reflections with $l=18$–20 remain visible on the pattern of the single-phase mixtures. These reflections correspond approximately to the so-called backbone spacing of the paraffin chain, 1·27A. The presence of these reflections indicates that thermal vibrations are less violent in the direction of the chains than normal to them.

The hexagonal structure of the single-phase mixtures is more symmetrical than that of the pure ketones and the volume of its unit cell is about

3 per cent larger than for a pure ketone. The hexagonal form thus represents a disordered packing of paraffin chains. It is analogous to the hexagonal high temperature form which occurs in some paraffins (Müller 1932). Müller explained this structure as being due to a twisting and turning of the chains which makes their cross section statistically circular instead of flat. It is interesting that the single-phase form of the mixed ketones is hexagonal over a wide temperature range, while the paraffin of the same chainlength melts without becoming hexagonal. This may be due to the presence of ketogroups at different levels. The dipole moment of these groups implies that they repel or attract each other, according to their configuration, and this is likely to cause some twisting of the chains even in their state of lowest potential energy.

On the whole, the X-ray data of the mixtures of equivalent ketones are in good agreement with the theory. They show that the single-phase form is more symmetrical and thus less ordered than the multiphase form, being hexagonal, while the multiphase form is orthorhombic or monoclinic. The multiphase form has a structure which is consistent with a physical mixture of the pure ketones. However, it has to be borne in mind that the X-ray data could not establish the presence or absence of dipolar planes. It would be desirable to refine the X-ray investigations of these mixtures, using single crystals.

Observations Under the Microscope.

Although the mixtures of equivalent ketones were not available as single crystals, for the purpose of the X-ray work, they could be made to grow crystallites distinct enough for investigation under the microscope. The technique used was as follows*:—a small quantity of mixture, of the order 5–10 milligrams, is put on a glass slide and covered with a coverslip. The mixture is then molten and the weight of the coverslip makes it spread in a thin film. If the specimen is now cooled slowly, only one or a few crystals form in the direction of the thickness. The crystals tend to be oriented with their *c* axes normal to the glass plates, the more so the slower the cooling. In any specimen it is generally possible to find some crystals filling most of the thickness and with their *c* axes normal to the plates.

Observation in convergent, polarized light, shows that all single phase mixtures are optically uniaxial, the *c* axis being the unique axis. This is in agreement with the hexagonal structure deduced from the X-ray diffractions pattern. The multi-phase mixtures are optically biaxial, with a small axial angle, and no significant difference can be found between these mixtures and the pure ketones. This agrees with their X-ray diffraction pattern.

---

* The author wishes to thank Mr. Scott Harley of the Burmah Oil Company, Ltd., for pointing out this method.

In parallel polarized light metastable mixtures of four ketones may be observed to change into the stable form. The hexagonal form presents a picture in which dark areas—crystals with their unique c axes normal to the plate—alternate with coloured crystals of other orientations (see Pl. XXI.(a)). When the hexagonal form changes into the orthorhombic form this picture changes. A pattern appears within the dark areas, which consists of bright birefringent needles (see Pl. XXI.(b)). These needles seem to represent minute cracks, more or less in crystallographic directions. The contraction manifested by these cracks is in agreement with the fact that the volume of the hexagonal form is 3 per cent larger than that of the monoclinic form.

A Low Temperature Transition in the Single-Phase Mixtures.

When single-phase mixtures of ketones were observed under the microscope, it was found that at about 15° C. they change into a form very similar to the multi-phase form (see Pl. XXI.(c)). This change was observed in detail on a microscope stage which could be cooled, using solid CO_2 (details of this microscope stage will be described elsewhere). It was found that on cooling from room temperature birefringent cracks began to appear at about 14·5–15·5° C. and became more intense as the temperature fell, until their intensity reached saturation round 13·5–14° C. These cracks make a pattern very similar to that formed when a single phase mixture of four ketones changes into its multi-phase form. The only difference is that now the cracks are longer and more distinct. On heating, the cracks vanish again, but at a slightly higher temperature; they begin to close at about 14·5–15·5° C. and the last cracks disappear at about 16–17° C. The transformation is rapid, it follows the change of temperature of the microscope stage without observable delay.

The low temperature transition occurs equally in all single phase mixtures of four, five and six ketones, at the same temperatures. The crystal structure of this form was examined by X-rays using a cooling attachment to a " Unicam " single-crystal goniometer. This attachment will be described elsewhere. Pl. XX.(c) shows the X-ray powder pattern of the low temperature form. It is indistinguishable from the pattern of the multi-phase form, also with regard to the higher order lines not shown on the Plate. When the multi-phase form is cooled below 14° C. it remains unchanged, both with regard to its X-ray structure and its appearance under the microscope.

As for the physical significance of the low temperature transition, it might be thought at first sight that it corresponds to a segregation of the ketones into the multi-phase form, for equation (8) shows that at sufficiently low temperatures it is always the multi-phase form which is stable. If a mixture of five or six ketones is, in equilibrium, single-phase at its melting point, it must, in equilibrium, break up into the n-phase form

below a temperature given by equation (8). However, the low temperature transition cannot correspond to this effect, for two reasons. Firstly, it occurs at the same temperature, whether the mixture concerned consists of four, five or six ketones. This contradicts equation (8). Secondly, the transition is very rapid. This is in contradiction with the observation that it takes several days for a mixture of four ketones to change into the stable form, at 18–20° C. At 15° C. the time for the change should be still longer. The rapidity of the low temperature transition implies that it cannot be due to a lengthy process of diffusion.

As the low temperature transition is not due to a segregation of different ketones, it must be due to a contraction of the hexagonal form, the ketogroups remaining distributed at random. The existence of such a transition is not very surprising because it is known that pure paraffins show transitions of this kind from an almost hexagonal high temperature form to a monoclinic or orthorhombic low temperature form (Müller 1932). A mixture of ketones without dipolar planes should not be very different from a paraffin if we assume that the chief contribution of the ketogroup to the binding of the structure is their dipolar interaction. Indeed, the single phase mixtures of five and six equivalent ketones have melting points of 27·5 and 27·0 respectively, as compared with 28° C. for the paraffin $C_{18}H_{38}$. However, the comparison is complicated by the fact that the paraffin $C_{18}H_{38}$ has a structure in which the chains are inclined to the basal plane, while they are normal to it for the mixtures of ketones.

Table II.*b* shows the dimensions of the unit cell for the low temperature form of the single-phase mixtures. These agree with the cell dimensions of the multi-phase form, within the limits of error. As far as the intensities are concerned the powder patterns of the single-phase low temperature form and the multi-phase form are indistinguishable. This implies that the packing of the chains is very similar.

The information provided by the X-ray data makes it possible to derive an expression for the free energy of the low temperature single-phase form. The free energy of the hexagonal form (equation (5)) contains an entropy of mixing $R \log 2n$. This should be the same for the low temperature form, if the ketogroups remain distributed at random between the possible levels. However, there should be no further difference between the entropy of the low temperature form and that of the pure ketone, because the volume and arrangement of the paraffin chains is the same for both. This should imply an equal mobility of the paraffin chains in the first approximation. Thus the entropy term for the low temperature form is $S_0+R \log 2n$. As for the total energy, it should differ from that of a pure ketone only in that dipolar planes are not present. Thus the total energy of the low temperature form should be H_0+h_D, and the free energy

$$G_3=H_0+h_D-T(S_0+R \log 2n). \quad . \quad . \quad . \quad . \quad . \quad (9)$$

The transition from the low temperature form to the hexagonal form may in principle be a transition of the first or second order. The experimental evidence indicates that it occurs over a temperature interval of 1–2° C. and is thus rather sharp. (The difference between the temperature of the transition or heating and cooling appears to be a hysteresis effect). Thus it seems justifiable to treat the transition as one of the first order. If this is done the transition temperature, T_3, can be derived by equating the free energies of the low temperature and hexagonal forms, which gives

$$h_x = T_3 s. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

The Energy of a Dipolar Plane.

Comparison of equation (10) with with equation (8) shows that the two equations contain three unknown variables, namely h_x, h_D and s. Attempts have been made to find a third equation containing these variables, so that they could be determined, but so far without success. However, equations (8) and (10) allow and approximate calculation of h_D, provided some rough estimate may be found for s, because

$$h_D = RT_1 \log 2n + (T_1 - T_3)s. \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

Here $n=4{\cdot}6$, T_1 is the transition temperature between multi- and single-phase mixtures for $n=4{\cdot}6$, namely 301° K, and T_3 is the transition temperature between hexagonal and low temperature forms, namely 288° K. As $T_1 - T_3$ is small, h_D is little sensitive to the value of s.

It is easy to find an upper limit for s, because it must certainly be less than the entropy of melting for a pure ketone. This has been measured by Oldham and Ubbelohde (1939) for the series of equivalent ketones of chainlength seventeen, and found to be 41·7 cal/mole degree on the average (see Table I.). This value for s gives $h_D < 1310 + 550$ cal/mole. However, the entropy of melting is certain to be higher than the entropy of a transition in the solid, and it is possible to find a closer approximation for s from the data of Ubbelohde (1938) or Garner, van Bibber and King (1931) who measured specific heats and latent heats of transitions for different paraffins. Ubbelohde's data, which have been evaluated by Fröhlich (1949) give a change of entropy, connected with the transition, of about 6 cal/mole degree. Garner's data give an entropy of transition of about 21 cal/mole degree. Assuming a value intermediate between the two alternatives gives

$$h_D = 1500 \text{ cal/mole}$$

which should be a fair first approximation.

The experimental value for h_D may be compared with the energy of electrostatic interaction for a dipolar plane, which has been calculated by Fröhlich (1946) to be

$$\frac{U}{2} = -\frac{1{\cdot}75}{\epsilon_\infty}\frac{\mu^2}{r^3},$$

where μ is the dipole moment of the C=O group, r the distance between dipoles in the dipolar plane and ϵ_∞ the dielectric constant of the medium between them. The latter is rather an indefinite quantity because the dipoles are not far enough apart to permit the material between them to be considered as a continuum. This means that ϵ_∞ might have any value between unity and the non-polar part of the dielectric constant of a ketone, namely 2—2·5. Evaluation of U/2 with ϵ_∞ indeterminate gives

$$\frac{U}{2} = \frac{2080}{\epsilon_\infty} \text{ cal/mole},$$

and this may be compared with the experimental value $h_D = U/2 = 1500$ cal/mole. The two quantities become equal if $\epsilon_\infty = 1{\cdot}4$, which seems a very plausible value. Thus the experimentally determined energy of the dipolar planes can be accounted for on the basis of dipolar interaction.

Conclusions.

Ketones of the formula $C_pH_{2p+1}COC_qH_{2q+1}$ with $p \geqslant 3$, $q \geqslant 3$ and $p \neq q$ were considered as equivalent ketones throughout the present paper. This corresponds to the assumption that all the bonding forces of these molecules are equal except for the presence of a ketogroup at different parts of the paraffin chain. The ketogroup was assumed to influence the bonding only by means of its dipole moment. With these assumptions, the treatment of mixtures of equal amounts of equivalent ketones could be based exclusively on statistical mechanics and classical electrostatics. This model was on the whole surprisingly successful. It was found that, as far as the packing of the paraffin chains is concerned, the mixtures always have one of two structures, of which one is disordered compared with the other one. The conditions of stability of these structures could be derived from the theoretical model, and were found to depend on the configuration of the ketogroups. Two kinds of order–disorder transitions were observed, one corresponding to the closing of a multiple solid-solubility gap and the other analogous to an order–disorder transition in a paraffin.

In considering the results obtained it has to be borne in mind that the crystallographic data available are somewhat crude and need refining by means of single-crystal work. A more detailed crystallographic investigation would presumably reveal minor differences between mixtures, the structures of which were classed as identical in the present paper. Such an investigation might prove interesting from the point of view of X-ray optics. Apart from this, an X-ray study of the transition from a metastable single-phase to a stable multi-phase mixture might provide interesting data on internal diffusion.

The data obtained in the present work on the conditions of the order–disorder transitions permitted the calculation of the electrostatic energy of a dipolar plane, in first approximation, but not the calculation of the

other entropies and energies involved in the transitions. It would be interesting to determine those, for instance by the measurement of specific heats. Such measurement would also provide a check on the validity of the assumptions made.

Apart from their intrinsic interest the mixtures of equivalent ketones may provide a tool for studying the dependence of physical properties on structural features in long-chain materials. The different equivalent ketones on the one hand and the various mixtures of ketones on the other hand represent a number of closely related structures, in which structural features vary in different respects. This may prove useful, for instance, for the study of molecular vibrations.

Acknowledgments.

The author wishes to thank Mr. C. G. Garton for valuable suggestions, Dr. J. W. H. Oldham for the loan of samples of ketones and advice on chemical synthesis, and Professor H. Fröhlich and Dr. H. N. V. Temperley for helpful discussions on theoretical aspects of this paper. The author is indebted to the Director of the British Electrical and Allied Industries Research Association for his encouragement and for permission to publish this paper.

References.

Fröhlich, H., 1946, *Proc. Roy. Soc.* A, **185,** 399 ; 1949, *Theory of Dielectrics* (Oxford), p. 148.

Fowler, R., and Guggenheim, E. A., 1939, *Statistical Thermodynamics* (Cambridge).

Garner, W. E., van Bibber, K., and King, A. M., 1931, *J. Chem. Soc.*, 1533.

Jackson, W., 1935 a, *Proc. Roy. Soc.* A, **150,** 197 ; 1935 b, *Ibid.*, **153,** 158.

McArthur, I., 1944, *Proc. Leeds Phil. Soc.* **4,** 170.

Müller, A., 1928, *Proc. Roy. Soc.* A, **120,** 437 ; 1930, *Ibid,* **127,** 417 ; 1932, *Ibid.*, **138,** 514 ; 1937, *Ibid.*, **158,** 403 ; 1938, *Ibid.*, **166,** 316.

Oldham, J. W. H., and Ubbelohde, A. R., 1939, *Trans. Faraday Soc.*, **35,** 328 ; 1940, *Proc. Roy. Soc.* A, **176,** 50.

Piper, S. H., Chibnall, A. C., Hopkins, S. J., Pollard, A., Smith, J. A. B., and Williams, E. F., 1931, *Biochem. J.*, **25,** 2087.

Pelmore, D. R., 1939, *Proc. Roy. Soc.* A, **172,** 502.

Sillars, R. W., 1939, *Proc. Roy. Soc.* A, **169,** 66.

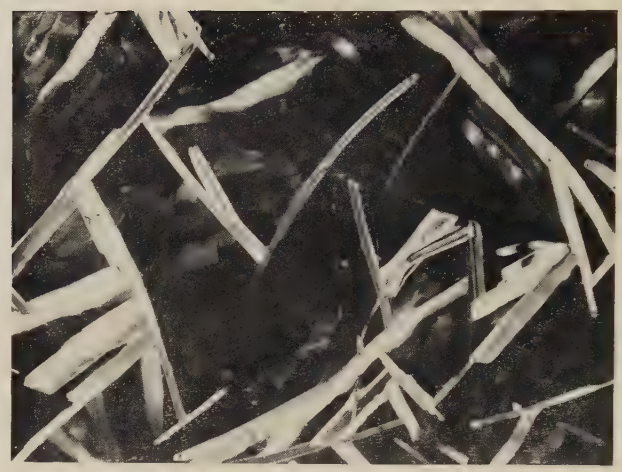
Ubbelohde, A. R., 1938, *Trans. Faraday Soc.*, **34,** 228.

(*a*)
$C_3+C_5+C_6+C_7$
metastable, 18° C.

(*b*)
$C_3+C_5+C_6+C_7$
stable, 18° C.

(*c*)
$C_3+C_5+C_6+C_7$
metastable, 11° C.

(*a*)
$C_3+C_4+C_5+C_7$
metastable, r.T.



(*b*)
$C_3+C_4+C_5+C_7$
stable, *r*.T.

(*c*)
$C_3+C_4+C_5+C_7$
metastable, 12° C.

LVIII. *On a Fourth-Order Meson-Equation.*

By W. Thirring,
Dublin Institute for Advanced Studies *.

[Received January 15, 1950.]

§ 1. Introduction.

It was shown recently by the author in a letter to the Editor of the *Physical Review* (Thirring 1950 a), that the quantization of higher order equations leads to results which correspond to the new subtraction method called regularization. But by applying the usual formalism of quantized fields to higher order equations one ought in general to interpret the field as belonging to particles which possess states of several rest masses. One is therefore confronted with the question whether one should accept a realistic or a formalistic interpretation. This question arises in the regularization method as well as in the field theories of Bhabha, Podolsky, etc. It is possible to circumvent this problem by using transcendental operators which possess only one zero-point in momentum space, as was done by Born and Green, but one encounters here the obstacle that a transcendental operator becomes hopelessly difficult, if one tries to introduce generalized derivatives in a space with an electromagnetic field. Another way of avoiding the problem of different rest masses is to take field equations of the form

$$(\Box^2-\mu^2)^n\phi(x)=0.$$

In this paper we shall discuss the next simplest case after the Klein–Gordon equation, namely

$$(\Box^2-\mu^2)(\Box^2-\mu^2)\phi(x)=0.$$

The time independent Green-functuon of this equation is the exponential potential

$$-\frac{\exp(-\mu r)}{2\mu}=\frac{1}{2\mu}\frac{\partial}{\partial\mu}\frac{\exp(-\mu r)}{r}.$$

As this potential has the advantage over the Yukawa potential that it makes no r^{-3} singularity in the tensor force, one is inclined to take this equation for the meson rather than the Klein–Gordon equation. By taking a reasonable meson mass and a simple scalar coupling we can adjust our coupling constant to well known nuclear data, and after having so fixed every constant we may proceed to calculate the self energy of the nucleon due to this meson field. It turns out, that the second order self energy is less than 8 per cent of the rest energy of the nucleus, but that in higher order approximations there are still divergent terms. The

* Communicated by the Author.

reason is that one can make the meson self energy due to virtual nucleon creation finite only by taking another nucleon equation but not by changing the meson equation. The goal of this paper is, of course, not to give a final solution of the divergence problem, but to show that by taking higher order equations some renormalization terms, which used to be infinite, are now finite and have an actual small numerical value. In this paper we use the new formalism of quantized fields, developed by Schwinger and the calculation method of Dyson–Feynman.

§ 2. Formalism.

First we shall define our notation, which is very similar to the one used by Feynman (1949). We choose natural units $\hbar=c=1$ and take real world coordinates with the metric tensor $g_{00}=1$, $g_{11}=g_{22}=g_{33}=-1$. The scalar product of four-vectors we write $a\,b=a_0b_0-\vec{a}\vec{b}$, an arrow indicates a spatial vector. The four-dimensional volume is written with out index $dk=dk_0d\vec{k}=dk_0\,dk_1\,dk_2\,dk_3$. By

$$\frac{\partial}{\partial x_\mu} \quad \text{we mean} \quad \frac{\partial}{\partial x_0},\ -\frac{\partial}{\partial x_1},\ -\frac{\partial}{\partial x_2},\ -\frac{\partial}{\partial x_3},$$

the d'Alembertian is

$$\Box^2=-\frac{\partial}{\partial x_\mu}\frac{\partial}{\partial x_\nu}g_{\mu\nu}.$$

The surface-element of a space-like surface is

$$d\sigma_\mu=-dx_1dx_2dx_3,\ dx_0dx_2dx_3,\ dx_0dx_1dx_3,\ dx_0dx_1dx_2.$$

The nucleons are represented by a quantized Dirac field

$$(i\nabla-m)\psi(x)=0, \qquad \nabla=\gamma_\mu\frac{\partial}{\partial x_\mu}=\beta\frac{\partial}{\partial t}+\beta\vec{\alpha}\,\text{grad}.$$

The γ's obey the commutation rules $\gamma_\mu\gamma_\nu+\gamma_\nu\gamma_\mu=2g_{\mu\nu}$. The scalar product of a four-vector with the matrix vector we write $\boldsymbol{\alpha}=a_\mu\gamma_\mu$.

The adjoint spinor is given by $\bar{\psi}=\psi^*\gamma_4$, $\bar{\psi}\underset{(\leftarrow)}{(i\nabla}+m)=0$. In momentum space the equations are :

$$\psi=U_p\exp(-ipx), \quad (\mathbf{p}-m)U_p=0, \quad \bar{U}_p(\mathbf{p}+m)=0, \quad p^2=m^2.$$

For the vacuum expectation value we have [Dyson (1949), Feynmen (1949)]

$$P\langle\psi(x),\ \bar{\psi}(x')\rangle_0=\tfrac{1}{2}\epsilon(x-x')(i\nabla+m)\Delta_F(x-x')$$

$$=\tfrac{1}{2}\epsilon(x-x')\frac{2i}{(2\pi)^4}\int\frac{\mathbf{p}+m}{p^2-m^2}\exp\{-ip(x-x')\}\,dp$$

$$\epsilon(x)=\frac{x_0}{|x_0|}.$$

P is the operator which arranges factors chronologically. We shall only consider free nucleons with sharp energy-momentum, and get therefore for the one particle expectation value

$$\mathrm{P}\langle\bar{\psi}(x),\quad \psi(x')\rangle_1=\epsilon(x-x')\bar{\mathrm{U}}_p\mathrm{U}_p\exp\{ip(x-x')\}.$$

Coming now to the meson field we first carry out quantization of the vacuum field and enter then immediately into the interaction representation for the field interacting with the nucleons. From here we can then deduce easily the field equations in the Heisenberg representation. Our vacuum equation is

$$(\Box^2-\mu^2)(\Box^2-\mu^2)\phi_i(x)=0. \qquad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

The index i indicates the component of the field function in isotopic spin space. Equation (1) is derived from a Lagrangian L

$$\mathrm{L}=-\frac{1}{2}\left(\frac{\partial^2}{\partial x_\varrho^2}\phi_i\frac{\partial^2}{\partial x_\sigma^2}\phi_i-2\mu^2\frac{\partial\phi_i}{\partial x_\varrho}\frac{\partial\phi_i}{\partial x_\varrho}+\mu^4\phi_i\phi_i\right). \qquad . \quad . \quad . \quad (2)$$

The commutation relations for a field with second order Lagrangian are given by Weiss (1938), Podolsky and Kikuchi (1944), Chang (1946), Chang (1948), De Wet (1948), Le Couteur (1948)

$$\int\left[\frac{\partial\phi_i(x)}{\partial x_0},\quad\frac{\partial\mathrm{L}\{x'\}}{\partial(\partial^2\phi_k(x')/\partial^2x_0')}\right]d\mathrm{V}'=i\delta_{ik}, \qquad . \quad . \quad . \quad . \quad (3)$$

$$\int\left[\phi_i(x),\quad\frac{\partial\mathrm{L}\{x'\}}{\partial(\partial\phi_k(x')/\partial x_0')}-\frac{\partial}{\partial x_j'}\frac{\partial\mathrm{L}\{x'\}}{\partial(\partial^2\phi_k(x')/\partial x_k'\partial x_j')}\right]d\mathrm{V}'=i\delta_{ik}. \qquad . \quad . \quad . \quad (4)$$

In our case, if we replace the space integration by the integration over an arbitrary space-like surface, we get

$$\int_\sigma\left[\frac{\partial\phi_i(x)}{\partial x_k},\quad\Box'^2\phi_j(x')\right]d\sigma_k=-i\delta_{ij}, \qquad . \quad . \quad . \quad . \quad (3a)$$

$$\int_\sigma\left[\phi_i(x),\quad 2\mu^2\frac{\partial}{\partial x_k'}\phi_j(x')-\frac{\partial}{\partial x_k'}\Box'^2\phi_j(x')\right]d\sigma_k=-i\delta_{ij}. \qquad . \quad . \quad (4a)$$

This gives for the commutator function

$$[\phi_i(x),\quad\phi_k(x')]=\delta_{ik}i\mathrm{D}(x-x') \qquad . \quad . \quad . \quad . \quad . \quad (5)$$

$$\int_\sigma\frac{\partial}{\partial x_\mu}\Box^2\mathrm{D}(x)\,d\sigma_\mu=-1,\qquad\int_\sigma\frac{\partial}{\partial x_\mu}\mathrm{D}(x)\,d\sigma_\mu=0. \qquad . \quad . \quad . \quad (6)$$

σ is here any space-like surface through the origin. $\mathrm{D}(x)$ has to be a solution of (1), therefore the associated function

$$\bar{\mathrm{D}}(x)=\tfrac{1}{2}\epsilon(x)\,\mathrm{D}(x), \qquad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

obeys (1) everywhere but at the origin. To investigate its behaviour there we consider the integral

$$\oint_k dx[\Box^2-\mu^2]^2\,\bar{\mathrm{D}}(x), \qquad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

where k is any region round the origin. Applying Gauss' theorem (8) becomes

$$\int_\sigma dS_l \frac{\partial}{\partial x_l}(\Box^2-2\mu^2)\overline{D}(x)+\int_k dx\,\mu^4\overline{D}(x)$$

σ is the surface of k.

If we make the region infinitesimally thin and choose the surface to be space-like we get

$$(8)=\int_\sigma d\sigma_l \frac{\partial}{\partial x_l}(\Box^2-2\mu^2)D(x)=-1.$$

$\overline{D}(x)$ is therefore the invariant Green-function of (1), the same connection with opposite sign holds between the Green-function and the function in the case of the second order field. The invariant Green-function of (1) is

$$\overline{D}(x)=\frac{-1}{(2\pi)^4}\int\frac{\exp(-ikx)}{(k^2-\mu^2)^2}dk=\frac{1}{2\mu}\frac{\partial}{\partial\mu}\overline{D}_\mu(x). \quad . \; . \; . \quad (9)$$

$\overline{D}_\mu$ is the Green-function of the Klein–Gordon equation. Though it is evident from the integral-representation one can easily prove that (9) is the right Green-function of (1)

$$(\Box^2-\mu^2)^2\overline{D}(x)=\frac{1}{2\mu}(\Box^2-\mu^2)\frac{\partial}{\partial\mu}(\Box^2-\mu^2)\overline{D}_\mu(x)+(\Box^2-\mu^2)\overline{D}_\mu(x)$$
$$=-\delta(x). \quad . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \quad (10)$$

For the commutator function we obtain

$$D(x)=2\epsilon(x)\overline{D}(x)=-\frac{1}{2\mu}\frac{\partial}{\partial\mu}D_\mu(x). \quad . \; . \; . \; . \; . \quad (11)$$

$D_\mu(x)$ is the D-function for the second order field with rest mass μ and as is necessary, (11) is an odd invariant solution of (1). For the calculation we need Feynman's function which is

$$P<\phi_i(x)\phi_k(x')>_0=\frac{\delta_{ik}}{2}D_F(x)=\frac{\delta_{ik}}{2}\frac{-2i}{(2\pi)^4}\int\frac{\exp(-ikx)}{(k^2-\mu^2)^2}.$$
$$. \; . \; . \quad (5a)$$

One notices the similarity between (5a) and Feynman's cut-off function.

Passing from the vacuum fields to interacting fields we keep our field equations, but the state vector is now no longer a constant but varies according to the fundamental equation of the interaction representation

$$\frac{\delta\Psi(\sigma)}{\delta\sigma(x)}=-i\mathfrak{H}(x)\Psi(\sigma). \quad . \; . \; . \; . \; . \; . \; . \quad (12)$$

For the interaction energy $\mathfrak{H}(x)$ we take a simple scalar coupling

$$\mathfrak{H}(x)=-\mu\sqrt{(2)}g_i\phi_i(x)\bar{\psi}(x)\tau_i\psi(x). \quad . \; . \; . \; . \; . \quad (13)$$

The factor μ is introduced to make the coupling constant g_l dimensionless, τ_l are the isotopic spin operators. The field equations in the Heisenberg representation are easily available. If we indicate the field quantities in the Heisenberg representation for a moment by Clarendon type, the general connection between the derivatives of the field quantitites in the Heisenberg representation and the interaction representation is given by Schwinger (1948)

$$\frac{\partial \mathbf{F}(x)}{\partial x_\nu} = \mathrm{U}^{-1}(\sigma)\frac{\partial \mathrm{F}(x)}{\partial x_\nu}\mathrm{U}(\sigma) + i\mathrm{U}^{-1}(\sigma)\int_\sigma [\mathrm{H}(x), \mathrm{F}(x')]\, d\sigma'_\nu\, \mathrm{U}(\sigma). \quad (14)$$

U is the unitary operator which performs the transformation between the Heisenberg and the interaction representations. Using (6), (13), (14) we obtain

$$\left.\begin{aligned} &\frac{\partial \boldsymbol{\varphi}_i}{\partial x_\nu} = \mathrm{U}^{-1}\frac{\partial \phi_i}{\partial x_\nu}\mathrm{U}, \qquad \Box^2\boldsymbol{\varphi}_i = \mathrm{U}^{-1}\Box^2\phi_i\mathrm{U}, \\ &\frac{\partial}{\partial x_\nu}\Box^2\boldsymbol{\varphi}_i = \mathrm{U}^{-1}\frac{\partial}{\partial x_\nu}\Box^2\phi_l\mathrm{U}, \\ &\Box^2\Box^2\boldsymbol{\varphi}_i = \mathrm{U}^{-1}\Box^2\Box^2\phi_i\mathrm{U} + \mathrm{U}^{-1}\mu\sqrt{(2)}g_i\bar{\psi}(x)\tau_i\psi(x)\mathrm{U}, \end{aligned}\right\} \quad . \quad . \quad . \quad (15)$$

therefore the field equations for the meson field in the Heisenberg representation are :

$$(\Box^2 - \mu^2)^2\boldsymbol{\varphi}_i = g_i\mu\sqrt{(2)}\bar{\boldsymbol{\psi}}\tau_i\boldsymbol{\psi}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (16)$$

To prove the third equation of (15) we need

$$\int_\sigma \Box^2\mathrm{D}(x)\, d\sigma_\nu = \frac{-1}{2\mu}\frac{\partial}{\partial\mu}\int_\sigma \mu^2 D_\mu(x)\, d\sigma_\nu = 0.$$

For the spinor field we get [Schwinger (1948)] if we use

$$\int\{\bar{\psi}_\alpha(x), \quad (\gamma_\mu\psi(x'))_\beta\}\, d\sigma'_\mu = \delta_{\alpha\beta} \quad \{\mathrm{A}, \mathrm{B}\} = \mathrm{AB} + \mathrm{BA},$$

and the identity

$$[\bar{\psi}(x)\psi(x), \psi(x')] = \psi(x)\{\bar{\psi}(x), \psi(x')\} - \{\bar{\psi}(x), \psi(x')\}\psi(x)$$

the following field-equations in the Heisenberg representation

$$(i\nabla - m + g_i\mu\sqrt{(2)}\tau_i\boldsymbol{\varphi}_i)\boldsymbol{\psi} = 0. \quad . \quad . \quad . \quad . \quad . \quad . \quad (17)$$

The field quantities in the Heisenberg representation are easily expressed in terms of the field quantities in the interaction representation [Schwinger (1949 b), Dyson (1949)]. For the interaction energy in the Heisenberg representation we get

$$\mathbf{H}(x_0) = \sum_{n=0}^{\infty}\frac{(-i)^n}{n!}\int_{-\infty}^{\infty} dx_1 \ldots \int_{-\infty}^{\infty} dx_n \mathrm{P}(\mathfrak{H}(x_0)\mathfrak{H}(x_1)\ldots\mathfrak{H}(x_n)). \quad (18)$$

First we consider the expectation value of the second term of the series (17) for a state where two nucleons and one meson are present, which is nothing but twice* the second order interaction between two nucleons.

* Half of it is subtracted if we transform H_0.

We get, if we integrate over the space volume,

$$2E_{\text{int}}=\int H_2(x_0)\, d\vec{x}_0$$
$$=-i2\mu g_k g_l \int d\vec{x}_0 dx_1 P<\psi(x_0)\tau_l\psi(x_0)\bar{\psi}(x_1)\tau_k\psi(x_1)\phi_l(x_0)\phi_k(x_1)>_{2,0}.$$

For simplicity we neglect, in evaluating the expectation value, exchange terms which give rise to mixed densities.

$$E_{\text{int}}=2\mu^2 g_l^2 \frac{1}{2(2\pi)^4}\frac{-1}{2\mu}\frac{\partial}{\partial\mu}\int_{-\infty}^{\infty}[\rho_l^1(x_0)\rho_l^2(x_1)+\rho_l^2(x_0)\rho_l^1(x_1)]$$
$$\times\frac{\exp\{-ik(x_0-x_1)\}}{k^2-\mu^2}d\vec{x}_0\, dx_1\, dk. \quad . \quad . \quad . \quad . \quad . \quad (19)$$

The upper index on the nucleon densities $\rho_l^i(x)=<\bar{\psi}(x)\tau_l\psi(x)>$ indicates to which nucleon the density belongs. To obtain the static energy we assume the densities to be time independent, carry out the time integration and get

$$E_{\text{int}}=-\mu\frac{g_l^2}{4\pi}\int\rho_l^1(\vec{x}_0)\rho_l^2(\vec{x}_1)\exp\{-\mu(x_0-x_1)\}\, d\vec{x}_0\, d\vec{x}_1. \quad . \quad . \quad (20)$$

Or if we assume the densities to be sufficiently concentrated and approximate them by a δ-function

$$E_{\text{int}}=-\mu\frac{g_l^2}{4\pi}\tau_l^1\tau_l^2\exp\{-\mu(x_1-x_2)\}. \quad . \quad . \quad . \quad . \quad (21)$$

x_1 and x_2 are the positions of the nucleons. The negative sign shows that we get attractive forces. The same result yields a classical meson theory if we start from equations (16) and (13) and replace the nucleon by a point source. We are now able to determine the coupling constant g. We may either choose symmetric theory $g=g_1=g_2=g_3$, $g_4=0$, or neutral theory $g=g_4$, $g_1=g_2=g_3=0$. In both cases $\mu g^2/4\pi$ is the strength of the potential. To fit the proton–proton scattering data for an exponential potential we need a strength of about 100 MeV. if we choose a meson mass of 282 m (Rosenfeld 1948). This gives a coupling constant $g^2/4\pi=0{\cdot}71$. It is clear that the expontial potential wants a stronger coupling than the Yukawa potential with its singularity. Now everything is ready for working out the self energy.

§ 3. Calculation of the Second Order Self Energy.

The second order self energy is given by the expectation value of the second term in (17) for a state where one nucleon and no meson is present. The expectation value

$$<\bar{\psi}(x_0)\psi(x_0)\bar{\psi}(x_1)\psi(x_1)\phi(x_0)\phi(x_1)>_{1,0}$$

is of course

$$<\bar{\psi}(x_0)\psi(x_0)\bar{\psi}(x_1)\psi(x_1)>_1<\phi(x_0)\phi(x_1)>_0.$$

The one-particle part is Schwinger (1949 a)

$$<\bar{\psi}(x_0)\psi(x_0)\bar{\psi}(x_1)\psi(x_1)>_1 = <\bar{\psi}(x_0)\psi(x_1)>_1 <\psi(x_0)\bar{\psi}(x_1)>_0 + <\psi(x_0)\bar{\psi}(x_1)>_1 <\bar{\psi}(x_0)\psi(x_1)>_0$$

which makes the self energy

$$\begin{aligned} \mathrm{E}_{\mathrm{self}} &= \int <\mathrm{H}_2(x_0)>_{1,0} d\vec{x}_0 \\ &= -i2\mu^2 g_l g_k \int d\vec{x}_0\, dx_1 <\tau^l_{\alpha\beta}\tau^k_{\gamma\delta}>_1 \mathrm{P}<\phi_l(x_0)\phi_k(x_1)>_0 \\ &\quad \times [\mathrm{P}<\bar{\psi}_\alpha(x_0)\psi_\delta(x_1)>_1 \mathrm{P}<\bar{\psi}_\gamma(x_1)\psi_\beta(x_0)>_0 \\ &\quad + \mathrm{P}<\bar{\psi}_\gamma(x_1)\psi_\beta(x_0)>_1 \mathrm{P}<\bar{\psi}_\alpha(x_0)\psi_\delta(x_1)>_0]. \quad . \ . \ . \quad (22) \end{aligned}$$

Inserting (5a) and the expectation values for the spinor field and taking into account that the vacuum expectation value of two spinors gives Kronecker operators in spin and isotopic spin spaces we get

$$\begin{aligned} \mathrm{E}_{\mathrm{self}} &= -i4\mu^2 g_l^2 \int d\vec{x}_0\, dx_1\, dp'\, dx\, \epsilon(x_0-x_1)\bar{\mathrm{U}}_p \exp\{ip(x_0-x_1)\} \\ &\quad \times \epsilon(x_0-x_1)\frac{i}{(2\pi)^4}\frac{\mathbf{p}'+m}{p'^2-m^2}\mathrm{U}_p \frac{-i}{(2\pi)^4}\frac{\exp\{-ik(x_0-x_1)\}}{(k^2-\mu^2)^2} <\tau_l\tau_l>_1. \\ &\quad . \ . \ . \quad (23) \end{aligned}$$

It is easy to see that the two different terms in the one particle expectation value give rise to the same integral (23) with different sign of k. The sign can be restored by a change of integration variable, we may therefore replace the second term by a factor two in the first one. Further, the square of every isotopic matrix τ is the unit matrix, all kinds of mesons give therefore the same contribution to the self energy, which is equal for neutron and proton. We will now drop the index on the coupling constant and calculate the contribution to the self energy of one kind of meson. The integration over space and nucleon momentum in (23) yields then

$$\mathrm{E} = -i\frac{4\mu^2 g^2}{(2\pi)^4}\int \bar{\mathrm{U}}_p \frac{\mathbf{p}-\mathbf{k}+m}{(k^2-\mu^2)^2(k^2-2pk)}\mathrm{U}_p\, dk. \quad . \ . \ . \quad (24)$$

We have integrated over the x_0 unit volume, because we shall renormalize our spinor function to the unit volume. To carry out the operation in spin space which is required in (24) we observe that it is compatible with the Dirac equation to choose the vector $\bar{\mathrm{U}}_\alpha(p)$ in spin space so that $\bar{\mathrm{U}}_\alpha \mathrm{U}_\beta = \mathrm{N}(\mathbf{p}+m)_{\beta\alpha}$. As desired this corresponds to an unpolarized wave since $\bar{\mathrm{U}}\sigma_{\mu\nu}\mathrm{U}=0$. The normalization constant N we may take to normalize either to one mesonic charge per unit volume $\bar{\mathrm{U}}_\alpha \mathrm{U}_\alpha = 1$, $\mathrm{N}=1/4m$ or to one electric charge per unit volume $\bar{\mathrm{U}}_\alpha \gamma^0_{\alpha\beta}\mathrm{U}_\beta = 1$, $\mathrm{N}=4p_0/1$. For convenience we take the former possibility.

After an easy spur computation the integral (24) is then evaluated by a straightforward calculation using Feynman's method (Feynman 1949) and we obtain

$$E_{self}=\frac{-g^2\mu^2}{2(2\pi)^2m^2}\int_0^1\frac{m2x(2-(1-x))}{(1-x)^2+x\mu^2/m^2}dx=-\frac{4g^2}{(4\pi)^2}m\left(\frac{\mu}{m}\right)^2\int_0^1\frac{x^2+x}{(x-x_1)(x-x_2)}dx. \quad . \; . \; . \quad (25)$$

The zeros of the denominator in (25) are worked out to be

$$x_{1,2}=\exp(\pm i\gamma),\quad \tan\gamma=\frac{\sqrt{(2\alpha-\alpha^2)}}{1-\alpha},\quad \alpha=\frac{\mu^2}{2m^2}=0{\cdot}0118,\quad \gamma=0{\cdot}153. \quad . \; . \; . \quad (26)$$

Introducing the further abbreviations

$$x_{12}-1=-\alpha\pm i\sqrt{(2\alpha-\alpha^2)}=\mathrm{P}\exp(\pm i\beta),\quad \mathrm{P}=0{\cdot}1535,\quad \beta=1{\cdot}647, \quad . \; . \; . \quad (27)$$

the elementary integration yields

$$\left.\begin{aligned}\int_0^1\frac{x^2}{(x-x_1)(x-x_2)}dx&=1+\frac{1}{\sin\gamma}\{\sin 2\gamma\,.\,\ln\mathrm{P}+(\beta-\gamma)\cos 2\gamma\}=6{\cdot}3\\ \int_0^1\frac{x}{(x-x_1)(x-x_2)}dx&=\frac{1}{\sin\gamma}\{\sin\gamma\ln\mathrm{P}+(\beta-\gamma)\cos\gamma\}=7{\cdot}7.\end{aligned}\right\}. \quad (28)$$

Inserting into (25) we get

$$E=\frac{-g^2}{\pi}\left(\frac{\mu}{m}\right)^2\frac{14{\cdot}0}{4\pi}m=-m\frac{g^2}{\pi}0{\cdot}025=-m\,0{\cdot}076. \quad . \; . \; . \quad (29)$$

The negative sign corresponds to the fact that we have attractive forces. For pseudoscalar coupling the self energy is even less, namely

$$E_{self(ps)}=-m\frac{f^2}{\pi}0{\cdot}00454.$$

The second order self energy gives therefore only a small contribution to the observable mass of the nucleon.

It might be interesting to calculate the classical self energy. If we replace the nucleon density in (16) for the static case

$$(\Delta+\mu^2)^2\phi(x)=\rho(x) \quad . \; . \; . \; . \; . \; . \; . \; . \quad (16a)$$

by a point source

$$\rho(x)=g\mu\sqrt{(2)}\delta(\vec{x})$$

and use the above mentioned solution of (16a) we get for the self energy

$$E_{self(cl)}=-\frac{1}{2}\int\rho(x)\phi(x)\,d\vec{x}=-\frac{1}{2}\int g\mu\sqrt{(2)}\,\delta(\vec{x})\frac{g\mu\sqrt{(2)}}{2\mu}\frac{\exp(-\mu r)}{4\pi}d\vec{x}$$

$$=\frac{-\mu^2g^2}{(2\pi)^4}\int 2\pi\delta(k_0)\frac{dx\,dk}{(k^2-\mu^2)^2}=\frac{-g^2}{4\pi}m\frac{\mu}{2m}=\frac{-g^2}{4\pi}0{\cdot}0768m. \quad . \quad (30)$$

Comparing the integral representation of the classical self energy given in (28) with (24) we see that the self energy of the quantized field in (24) corresponds to a spread out of the nucleon density. This causes the reinforcement of the denominator in (24) which cuts off the contributions of the high frequencies and makes the quantum theoretical self energy a bit smaller than the classical one.

§ 4. Higher Order Approximations.

In the fourth order contribution to the self energy there are terms which are even smaller than the second order self energy and which correspond to double emission of mesons. The Feynman graphs of these terms would be

Fig. 1.

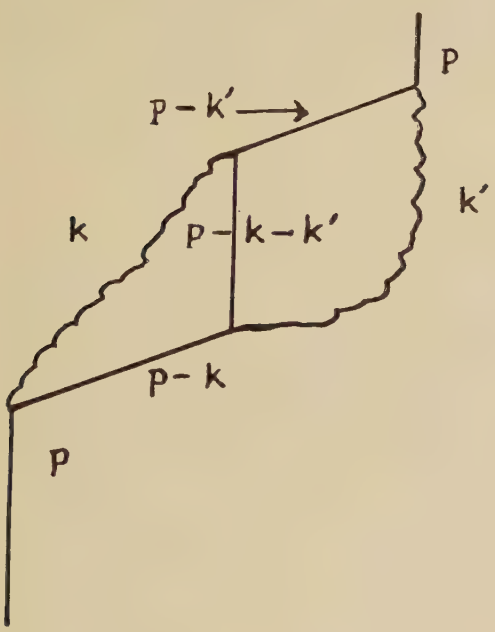
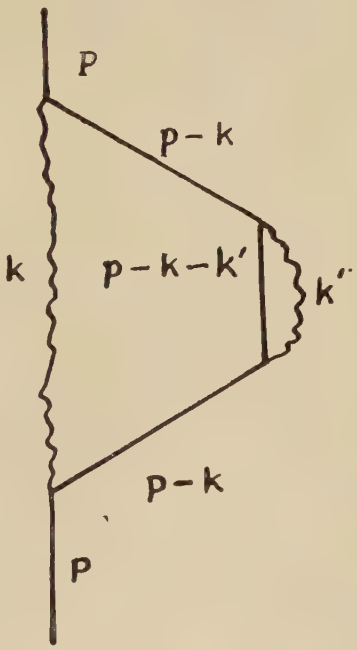


Fig. 2.



but there are still divergent terms arising from the creation of nucleons as in the following graph:

Calculating the contribution to the self energy due to this graph we encounter the following integral:

$$\int \frac{(\mathbf{p}-\mathbf{k}+m)tr(\mathbf{p}'+m)(\mathbf{p}'-\mathbf{k}-m)}{(k^2-\mu^2)^4(p'^2-m^2)((p'-k)^2-m^2)}\,dp'\,dk.$$

Though the denominator is of degree 12 and the numerator is of degree 11, this integral is divergent as it contains the self energy of the meson which has the divergent integral

$$\int \frac{tr(\mathbf{p}'+m)(\mathbf{p}'-\mathbf{k}+m)}{(p'^2-m^2)((p'-k)^2-m^2)}\,dp'.$$

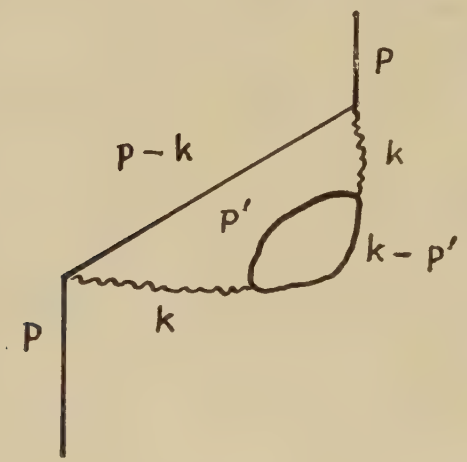
On this integral the form of the meson equation has no influence and it is infinite in our theory as well as in the usual one. The same fact mars the other higher order terms.

In conclusion we may state, that it is possible to make finite terms which are infinite in the usual theory, by taking higher order equations, but it does not yet seem to be possible to select from the manifold of linear equations of arbitrary order just those equations which might be realized in nature.

I should like to take this opportunity of thanking Professor Schrödinger for valuable discussions in connection with this paper.

Note added in proof.—In the meantime the same equation was proposed independently by Bhabha (*Phys. Rev.*, **77,** 665, 1950). For a closer discussion of the difficulties of this theory see Pais and Uhlenbeck (to appear in *Phys. Rev.*, **77,** 1st July, 1950).

Fig. 3.



References.

Chang, T., 1946, *Proc. Camb. Phil. Soc.*, **42,** 133 ; 1948, *Ibid.*, **44,** 76.
De Wet, J., 1948, *Proc. Camb. Phil. Soc.*, **44,** 546.
Dyson, F., 1949, *Phys. Rev.*, **75,** 486.
Feynman, P., 1949, *Phys. Rev.*, **76,** 769.
Le Couteur, K., 1948, *Proc. Camb. Phil. Soc.*, **44,** 63.
Podolsky, B., and Kikuchi, C., 1944, *Phys. Rev.*, **65,** 229.
Rosenfeld, L., 1948, *Nuclear Forces.* North-Holland Publishing Company, (Amsterdam).
Schwinger, J., 1948, *Phys. Rev.*, **74,** 1439 ; 1949 a, *Ibid.*, **75,** 651 ; 1949 b, *Ibid.*, **76,** 790.
Thirring, W., 1950 a, *Phys. Rev.*, **77,** 570 ; 1950 b, *Acta Phys. Austr.* (in the press).
Weiss, P., 1938, *Proc. Roy. Soc.* A, **169,** 102.

LIX. *The Effect of Electron Concentration on the Lattice Spacings in Magnesium Solid Solutions.*

By H. Jones,
Mathematics Department, Imperial College, London *.

[Received April 6, 1950.]

Summary.

The hypothesis that the change in the c/a ratio in magnesium solid solutions is due to the changes in the Fermi energy, arising from the distortion of the Brillouin zone, is developed into a more general and quantitative theory. The case of monovalent solutes is considered, and agreement with experiment is shown to be not unsatisfactory. The experimental results enable deductions to be made regarding the density of electronic states in pure magnesium.

§ 1. Introduction.

The changes which occur in the lattice parameters a and c of magnesium solid solutions, as the concentration of the solute is varied, have been measured with great accuracy by Hume-Rothery and Raynor (1940) and Raynor (1942). A simple qualitative theory has already been proposed to account for the variation of the c/a ratio in those solid solutions for which the valency of the solute is greater than that of magnesium (Jones 1949). The experimental results show unambiguously, in all cases, that this c/a ratio is determined, not by the atomic concentration of the solute, but by the valence electron concentration in the alloy.

It is the purpose of this paper to express the theory in a more general and quantitative form than hitherto, and to extend it to include those cases where the valency of the solute is less than the valency of magnesium, *e. g.* to the cases of lithium and silver.

The changes brought about in the overall form of the lattice by the addition of other elements in solid solutions may be regarded as the sum of two effects, (i) a change of atomic volume, and (ii) the development of a shearing strain. The change in the atomic volume must depend upon many factors, *e. g.* the difference between the size of the solute ion and the size of the magnesium ion, and also upon the difference in the potential fields of the solute and solvent ions in which the valence electrons move. The complexity of the causes of the change of volume can clearly be seen in the experimental observations which show no obvious regularities between the dilatation at a given concentration and the solute metal.

* Communicated by the Author.

Since the structure of pure magnesium is almost exactly that of hexagonally close packed spheres (the c/a ratio for Mg is 1·6237 and the value for ideal packing 1·6330), and since the metal is comparatively elastically isotropic, there is no reason to believe that differences of ionic size, or of spherically symmetrical potential fields, will play a significant part in producing the shearing strain. It might, therefore, be expected *a priori* that the shearing strain would result mainly from the effect of the Brillouin-zone structure on the valence electrons of the metal. This supposition is confirmed by the fact, already mentioned, that the shearing strain is substantially independent of the particular solute metal and depends, to a large degree, only on the ratio of the number of electrons to atoms in the metal.

§ 2. A Convenient Formulation of the Experimental Results.

It may be assumed that the relations between the lattice parameters and the concentrations of solute atoms will not be significantly different at the absolute zero of temperature from the observed relations at room temperature. Thus it will be sufficient to express the condition for equilibrium by minimizing the energy rather than the free energy.

The energy per atom W, expressed relative to the energy of the separated atoms, is a function of the atomic volume v, of the shearing strain, and, in the first place, of the solute concentration x. It is convenient to introduce the variable $s=(c/a)^{2/3}$ so that $a=(4v/\sqrt{3})^{1/3}s^{-1/2}$ and $c=(4v/\sqrt{3})^{1/3}s$. The conditions for equilibrium can then be written :

$$\text{(i)}\ \left(\frac{\partial \mathrm{W}(v,s,x)}{\partial v}\right)_{s,x}=0\ ; \qquad \text{(ii)}\ \left(\frac{\partial \mathrm{W}(v,s,x)}{\partial s}\right)_{v,x}=0. \quad . \quad . \quad (1)$$

In order to obtain the variation of c/a with x, as determined experimentally, we should have to eliminate v between these two equations. We shall make the assumption that the changes of volume arise from forces which act equally in all directions and can thus be simulated by the effect of a uniform pressure or tension. Such a uniform pressure or tension, in an anisotropic elastic crystal, produces a shearing strain given by

$$\frac{s-s_0}{s_0}=\tfrac{2}{3}k\left(\frac{v-v_0}{v_0}\right), \qquad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

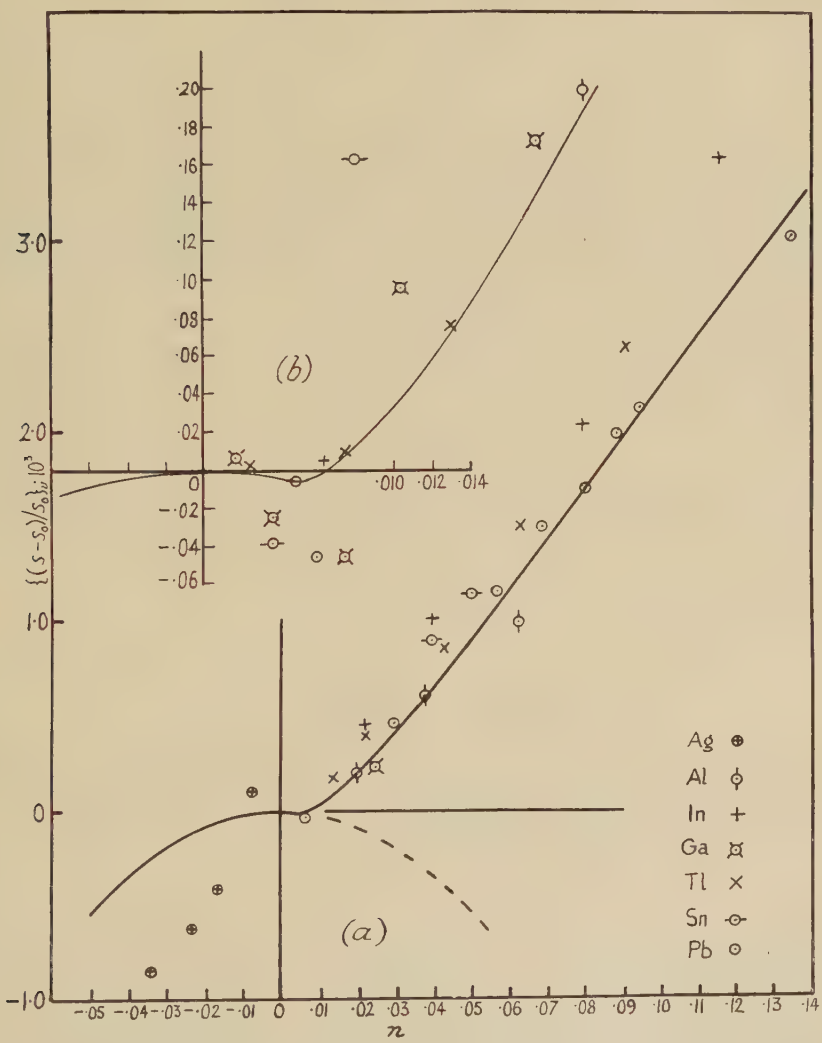
where

$$k=\frac{s_{13}+s_{33}-s_{11}-s_{12}}{2(s_{11}+s_{12})+4s_{13}+s_{33}}, \qquad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

and the s_{ij} are the elastic *coefficients* of the crystal. Thus, if we subtract this strain from the observed value, we get the strain induced by the solute atoms at *constant volume*. This quantity which we can determine from the experimental results can now be compared with the relation between s and x given by the second of equations (1) alone. The reason for this

procedure is that the dependence of W on v is too complicated to be amenable to calculation, whilst the dependence of s on x at constant v, we believe, arises from a particular effect which can be comparatively easily isolated.

Fig. 1.



(*a*) The variation of the c/a ratio with electron concentration in Mg solid solutions.
(*b*) Detail of (*a*) for very small concentrations.

If $\delta a/a$ and $\delta c/c$ are the observed strains, we get, for the change of s at constant volume,

$$\left(\frac{s-s_0}{s_0}\right)_v = \tfrac{2}{3}\left(\frac{\delta c}{c}-\frac{\delta a}{a}\right)-\tfrac{2}{3}k\left(\frac{\delta c}{c}+2\frac{\delta a}{a}\right)$$

$$= \tfrac{2}{3}(1-k)\frac{\delta c}{c}-\tfrac{2}{3}(1+2k)\frac{\delta a}{a}. \quad . \quad . \quad . \quad . \quad . \quad (4)$$

For pure magnesium $k=0{\cdot}0226$. In fig. 1 $\{(s-s_0)/s_0\}_v$ is plotted, not against atomic concentration, but against the increase of the number of electrons per atom in the alloy which we denote by n. If $\mathcal{N}$ is the number of valence electrons per atom of the solute, $n=(\mathcal{N}-2)x$. The values of $\mathcal{N}$ assumed for the solute metals are as follows:

	Li	Ag	Al	Ga	In	Tl	Sn	Pb
$\mathcal{N}=$	1	1	3	3	3	3	4	4

§ 3. A Theory of the Origin of the Shearing Strain.

Since fig. 1 shows clearly that variations of s at constant volume depend only on n, we shall use, from this point onwards, n as the variable in place of x; and also since all changes, now to be considered, take place at constant volume, there is no further need to make explicit reference to v, and therefore this quantity will no longer be shown as a variable. The condition for equilibrium becomes therefore $(\partial \mathrm{W}(s, n)/\partial s)_n=0$, which is the required equation to determine s as a function of n. For the derivative of s with respect to n, we have therefore

$$\frac{ds}{dn}=-\frac{\partial^2 \mathrm{W}}{\partial s \partial n}\Big/\left(\frac{\partial^2 \mathrm{W}}{\partial s^2}\right)_n. \qquad \ldots\ldots\ldots \quad (5)$$

Since (W/v) is equal to the strain energy, apart from a constant, and since the strain energy in terms of s can easily be shown to be $\frac{1}{2}\mu(s-s_0)^2/{s_0}^2$, where

$$\mu=\tfrac{1}{2}(c_{11}+c_{12}+2c_{33}-4c_{13}),$$

and c_{ij} are the elastic *constants* of the crystal, we find

$$\frac{1}{s_0}\frac{ds}{dn}=-\left(\frac{s_0}{\mu v}\right)\frac{\partial^2 \mathrm{W}}{\partial n \partial s}. \qquad \ldots\ldots\ldots \quad (6)$$

It will now be assumed that the only part of W which contributes to the second derivative on the right-hand side of (6) is the Fermi energy. This Fermi energy per atom, which we denote by w_{F}, may be considered as made up of several parts as follows:

$$w_{\mathrm{F}}=w_0+\Sigma w_i-\Sigma w_j, \qquad \ldots\ldots\ldots \quad (7)$$

where w_0 is the energy corresponding to the completely filled Brillouin zone of two electrons per atom. In magnesium or its alloys some electrons must overlap into the second zone, at various points, leaving vacancies in the first zone. The overlapping electrons and the holes may lie in distinct regions of momentum space. The energy of one such region containing n_i overlapping electrons is denoted by w_i and the energy of n_j electrons which would just fill one particular region of holes is written w_j. Thus we see that w_{F} is given by (7), where the summations are over all distinct regions of holes (j) and overlapping electrons (i).

w_i is a function of s and n_i only, but n_i is itself a function of s and n, and therefore

$$\left(\frac{\partial w_i}{\partial s}\right)_n=\left(\frac{\partial w_i}{\partial s}\right)_{n_i}+\left(\frac{\partial w_i}{\partial n_i}\right)_s\left(\frac{\partial n_i}{\partial s}\right)_n. \qquad \ldots\ldots \quad (8)$$

Now $(\partial w_i/\partial n_i)_s = \mathrm{E}_i(s, n_i)$, where E_i is the energy at the top of the filled levels in the region (i). For equilibrium, we have

$$\mathrm{E}_1 = \mathrm{E}_2 = \ldots = \mathrm{E}_i = \ldots = \mathrm{E}_j = \zeta, \quad \ldots \quad (9)$$

where ζ is the energy at the surface of the filled states, and thus

$$\Sigma \mathrm{E}_i \left(\frac{\partial n_i}{\partial s}\right)_n - \Sigma \mathrm{E}_j \left(\frac{\partial n_j}{\partial s}\right)_n = 0, \quad \ldots \quad (10)$$

since

$$n = \Sigma n_i - \Sigma n_j. \quad \ldots \quad (11)$$

Thus, since w_0 is independent of n, we have, using (8) and (10),

$$\frac{\partial^2 w_\mathrm{F}}{\partial n \partial s} = \Sigma \frac{\partial^2 w_i}{\partial s \partial n_i}\left(\frac{\partial n_i}{\partial n}\right)_s - \Sigma \frac{\partial^2 w_j}{\partial s \partial n_j}\left(\frac{\partial n_j}{\partial n}\right)_s. \quad \ldots \quad (12)$$

If $\mathrm{N}(\zeta)$ denotes the total density of states at the energy ζ, and $\mathrm{N}_i(\zeta)$ the contribution from the region of momentum space (i), then clearly

$$\left(\frac{\partial n_i}{\partial n}\right)_s = \frac{\mathrm{N}_i(\zeta)}{\mathrm{N}(\zeta)}, \quad \ldots \quad (13)$$

with exactly similar relations, apart from a negative sign, holding for the regions of the holes.

Thus (12) becomes

$$\frac{\partial^2 w_\mathrm{F}}{\partial n \partial s} = \frac{1}{\mathrm{N}(\zeta)}\left\{\Sigma \mathrm{N}_i \left(\frac{\partial \mathrm{E}_i}{\partial s}\right)_{n_i} + \Sigma \mathrm{N}_j \left(\frac{\partial \mathrm{E}_j}{\partial s}\right)_{n_j}\right\}. \quad \ldots \quad (14)$$

Finally, if we define coefficients ϕ_i by the equations

$$\phi_i = \frac{s_0}{\zeta}\left(\frac{\partial \mathrm{E}_i}{\partial s}\right)_{n_i}, \quad \ldots \quad (15)$$

(6) and (14) give

$$\frac{1}{s_0}\frac{ds}{dn} = -\left(\frac{\zeta}{v\mu}\right)\frac{1}{\mathrm{N}(\zeta)}\{\Sigma \phi_i \mathrm{N}_i + \Sigma \phi_j \mathrm{N}_j\}. \quad \ldots \quad (16)$$

§ 4. Comparison with Experimental Results.

In order to apply equation (16) to the magnesium solid solutions, we have to make detailed assumptions with regard to where the holes and the overlapping electrons occur in relation to the Brillouin zone. Fig. 2 (a) shows the form of the zone which contains, for all values of s, exactly two electrons per atom. We shall assume that the holes in the first zone occur at points along the line $\mathrm{D}_1\mathrm{D}_2$ and regard the centre of this line D as the point of highest energy *. We shall also assume that electrons overlap at the points A, and that for $n > 0$ overlap may also occur at the points B.

All points like those indicated by D in fig. 2 (a) are equivalent and lie at equal distances from the origin for all values of s. The neighbourhoods of D form altogether six identical regions in which the surfaces of

* Until calculations of the regions of highest energy within the Brillouin zone have been made any assumption such as this is necessarily arbitrary since the geometry of the zone is not sufficient to indicate the relative energies. It is of some interest that the coefficient ϕ_D depends upon the assumed position of the point of highest energy and that the experimental data relating to the variation of c/a are most conveniently explained by the present choice.

constant energy are simply connected closed surfaces. There are two separate regions corresponding to the points A, and one to the points B. The distances from the origin, k_A, k_B and k_D, are given by :

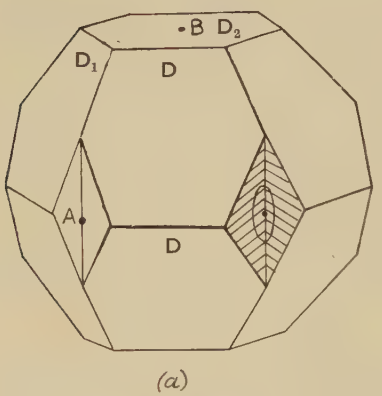
$$k_A=(\sqrt{3}/4v)^{1/3}\tfrac{2}{3}s^{1/2},$$

$$k_B=(\sqrt{3}/4v)^{1/3}\frac{1}{s},$$

$$k_D=(\sqrt{3}/4v)^{1/3}\frac{1}{\sqrt{3}}\left(1+\frac{3}{4s^3}\right)s^{1/2}.$$

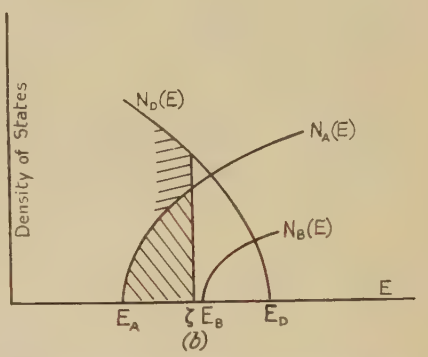
In order to calculate the coefficients ϕ_i, we notice, first, that $\partial E_i/\partial s$ is taken at constant n_i and, therefore, the changes in E_i must be closely equal to the changes in E_i when $n_i=0$, *i. e.* of the energy of the centre point of the region, and secondly, in equilibrium the value of E_i is always equal to ζ.

Fig. 2.



(*a*)

(*a*) Brillouin zone for Mg.



(*b*)

(*b*) Density of states for overlapping electrons N_A and for holes N_D.

For these reasons, we may assume that ϕ_A, for example, will be given without serious error by $d\log E_A/d\log s$. It is to be noticed that the ϕ's are unlikely to depend at all sensitively on n, and, of course, according to the above assumption, they are completely independent of n. For the application of (16), this is their most important property since it means that the right-hand side of this equation only varies with n significantly through the densities of states N_i. Numerical estimates of the ϕ's may be obtained by assuming that E_A, E_B and E_D vary as the squares of the distances from the origin to the points A, B and D respectively. In this way we find :

$$\phi_A=1, \qquad \phi_B=-2, \qquad \phi_D=1-\frac{18}{3+4s^3}=-\frac{13}{41}.$$

Thus (16) may be written

$$\frac{1}{s_0}\frac{ds}{dn}=\left(\frac{\zeta}{v\mu}\right)\frac{1}{N}\left\{2N_B+\frac{78}{41}N_D-2N_A\right\}, \qquad \ldots \quad (17)$$

where N_A, N_B and N_D denote the densities of states of *single* regions about their respective points.

Since, even for $n=0$, there is overlap at A and holes at D, we can expand N_A and N_D about their values at $n=0$. Thus, for instance,

$$N_A(n)=N_A^{(0)}+\left(\frac{dN_A}{dE}\right)_0\left(\frac{dE}{dn}\right)_0 n$$

$$=N_A^{(0)}+\frac{1}{N}\frac{dN_A}{d\zeta}n, \quad \ldots\ldots\ldots \quad (18)$$

where N is the total density of states in pure magnesium. N_B cannot, of course, be expanded in this way, since it is zero up to a certain (very small) value of n which we denote by n_0. For $n>n_0$, we may assume that

$$N_B=\alpha\sqrt{(\zeta-E_B)},$$

where α is a constant. Also we have $n-n_0=(\zeta-E_B)N$, and therefore

$$N_B=\frac{\alpha}{\sqrt{N}}(n-n_0)^{1/2}. \quad \ldots\ldots\ldots \quad (19)$$

Finally, substituting for N_A, N_B and N_D into equation (17), and integrating with respect to n, we find

$$\left(\frac{s-s_0}{s_0}\right)_v=A(n-n_0)^{3/2}+Bn+Cn^2, \quad \ldots\ldots \quad (20)$$

where

$$A=\left\{\begin{array}{ll}\frac{4}{3}\left(\frac{\zeta}{\mu v}\right)\frac{\alpha}{N^{3/2}}, & \text{if } n>n_0,\\ 0\quad, & \text{if } n<n_0,\end{array}\right\} \quad \ldots\ldots \quad (21)$$

$$B=\frac{2}{N}\left(\frac{39}{41}N_D^{(0)}-N_A^{(0)}\right)\left(\frac{\zeta}{\mu v}\right), \quad \ldots\ldots\ldots \quad (22)$$

and

$$C=\frac{1}{N^2}\left(\frac{39}{41}\frac{dN_D}{d\zeta}-\frac{dN_A}{d\zeta}\right)\left(\frac{\zeta}{\mu v}\right). \quad \ldots\ldots\ldots \quad (23)$$

The experimental results, for solutes with $\mathcal{N}>2$, indicate (*cf.* fig. 1) that B must be zero or exceedingly small. Hence we deduce that $N_A^{(0)}$ and $N_D^{(0)}$ must be very nearly equal.

It is clear from equation (23) that C must be negative, for $dN_D/d\zeta$ is necessarily negative and $dN_A/d\zeta$ positive (see *e. g.* fig. 2 (*b*)). Thus monovalent solutes, for which $n<0$, must decrease the c/a ratio. This conclusion is in agreement with the experimental results for lithium and silver.

It is possible to make a direct estimate of the number C by assuming that N_A and N_D have approximately parabolic form as shown in fig. 2 (*b*). If ϵ_1 is the energy width of the electrons overlapping at A, and ϵ_2 the energy width of the distribution of holes, then

$$\frac{dN_A}{d\zeta}=\frac{N_A}{2\epsilon_1}, \qquad \frac{dN_D}{d\zeta}=-\frac{N_D}{2\epsilon_2}.$$

For magnesium $\zeta=7{\cdot}3$ eV. as given by the soft X-ray emission spectra (Skinner 1939), $\mu=6{\cdot}23\times10^{11}$ dynes/cm.^{2}, and $v=23{\cdot}08\times10^{-24}$ cm.^{3} Thus the pure number $(\zeta/\mu v)=0{\cdot}813$. Hence, neglecting the difference between 39/41 and unity, we obtain

$$C=-\frac{0{\cdot}813}{2N^2}\left(\frac{N_A}{\epsilon_1}+\frac{N_D}{\epsilon_2}\right),$$

and, since $N=6N_D+2N_A$, and $N_A=N_D$, we get, if we assume that N may be taken as $3/\zeta$, which is the free electron value and probably slightly too big,

$$C=-\frac{0{\cdot}813\zeta}{48}\left(\frac{1}{\epsilon_1}+\frac{1}{\epsilon_2}\right).$$

If we take $\epsilon_1=\epsilon_2=1$ eV. which is roughly the value of ϵ_1 suggested by the X-ray data, we find $C=-0{\cdot}247$. The value of C which has been used to plot the curves of fig. 1 is $-0{\cdot}213$, and is thus quite compatible with the above estimate. The curves of fig. 1 are given by

$$\left(\frac{s-s_0}{s_0}\right)_v=-0{\cdot}213n^2+0{\cdot}150(n-0{\cdot}005)^{3/2}, \qquad n>0{\cdot}005,$$

$$=-0{\cdot}213n^2 \qquad , \qquad n<0{\cdot}005,$$

where it has been assumed that overlap sets in at the points B when n exceeds 0·5 per cent.

The agreement is not good in the case of silver, but the curve for negative n goes exactly through the only known point for the magnesium lithium solid solution, viz. that at the phase boundary limit of 17 atomic per cent lithium *. At this concentration the observed $(s-s_0)/s_0$ is equal to $-{\cdot}0067$.

It must be admitted that the interpretation of the experimental results, which we have adopted, depends largely on the observations at very small n. Apart from this region, a fairly good straight line could be drawn through the whole set of points both for positive and negative n. The interpretation would then be that the constants A and C were relatively small, and that B was positive, implying that $N_D>N_A$. To distinguish with certainty between these possibilities more data relating to the strain produced by monovalent solutes would be necessary.

References.

Hume-Rothery, W., and Raynor, G. V., 1940, *Proc. Roy. Soc.*, **177**, 27.
Jones, H., 1949, *Physica*, **15**, 13.
Raynor, G. V., 1940, *Proc. Roy. Soc.*, **174**, 457 ; 1942, *Ibid.*, **180**, 107.
Skinner, H. W. B., 1939, *Rep. Prog. Phys.*, **5**, 257 (London: Physical Society).

* The author is indebted to Professor G. V. Raynor for the information relating to the Mg-L_i system.

LX. *On the Correlation of the Directional Properties of Rolled Shee in Tension and Cupping Tests.*

By L. Bourne, B.Met., and R. Hill, M.A., Ph.D.,
Metal Flow Research Laboratory, Sheffield*.

[Received April 12, 1950.]

Summary

Plastic anisotropy in rolled sheet is examined by tension and cupping tests. The earing positions are correlated with the strain-ratios measured in tension tests at various orientations to the direction of rolling. The materials used are copper giving four ears at 45°, brass giving four ears at 50°, and brass giving six ears at 0° and 60°. A theory of plastic anisotropy due to Hill is found to be in good agreement with the experimental data for materials producing four ears. The theory is extended to describe more complex states of anisotropy.

§ 1. Method of Correlation.

The deformation imparted during the rolling of sheet metal produces an elongation of the crystal grains and a tendency towards a preferred orientation. The resulting anisotropy can be reduced or accentuated at will by a suitable programme of rolling reductions and interstage heat-treatment. Directionality in the final product is revealed by earing in a cupping test on a circular blank, or by the variation of mechanical properties in tension tests on specimens cut from the sheet at different angles to the direction of rolling. It is evident that the plastic behaviour of a metal with crystallographic anisotropy is directly dependent on the plastic properties of a single crystal of the metal and on the degree of preferred orientation. If the mechanism of distortion of a grain in an aggregate were completely known, it would be possible, in principle, to predict the behaviour of the aggregate when loaded in any way. This cannot yet be done since it is not known, for instance, how the operative glide system in an individual grain is influenced by the constraint of neighbouring grains. For example, in a face-centred cubic metal there are twelve operable slip-directions at ordinary temperatures, whereas a general uniform strain can be produced by a combination of shears in only five independent directions. Sufficient conditions for the determination of a unique active system are unknown. For this reason alone, preliminary attempts to relate the earing positions to the rolling texture seem to be premature†.

* Communicated by the Authors.

† For a general review of past work see T. Ll. Richards, " Progress in Metal Physics ", **1** (1949), pp. 281–305. (Butterworth's Scientific Publications, London.)

A more practical alternative is to measure the macroscopic properties of the anisotropic sheet in simple tests and, from the data so obtained, attempt to predict the behaviour of the sheet when deformed by any complex stress system. The question immediately arises as to which properties of the sheet are relevant to its behaviour under combined stress, in particular when deep-drawn into a cup. It is obvious that only *mechanical* properties are relevant, and that of these the breaking stress and elongation can be ignored when we are not concerned with fracture. Again, when we are interested only in *large* plastic deformations, the elastic moduli are irrelevant except in so far as they affect the residual stresses after unloading. There remains only the plastic stress–strain properties, which incorporate (i) the dependence of work-hardening on the amount of deformation, and (ii) the relations between the ratios of the components of the stress and the plastic strain-increment. In general, both factors govern the distortion of the sheet when taken along some complex loading-path. In the cupping test, however, the conditions are fairly simple, each element of the rim being subjected to a circumferential strain of approximately the same amount if the material is not too anisotropic. Moreover, ear-formation presumably depends mainly (though not entirely) on conditions near the rim, and here the state of stress is essentially a pure circumferential compression. The blank-holder pressure is usually negligible compared with the yield stress, and hinders the thickening of the edge only slightly because of the elasticity of the apparatus. Hence factor (ii) appears likely to predominate in controlling the positions of the ears. The relevant fundamental test is therefore the measurement of strain-ratios in uniaxial compression at various angles in the plane of the sheet. For convenience it is obviously preferable to perform tension tests instead, although this involves the assumption that the strain-ratios are unaffected by microscopic internal stresses which exert different influences in tension and compression. This seems entirely reasonable since the final annealing should remove those stresses responsible for the Bauschinger effect and allied phenomena. The measurement of strain-ratios in tension tests is the basis of the present method of predicting earing positions. In essence, the method has been used previously by Baldwin, Howald and Ross (1945), but the details differ, as will be explained later.

§ 2. Theoretical Considerations.

According to a theory of plastic anisotropy proposed by Hill (1948), the relations between the ratios of the components of the stress and strain-increment are

$$\frac{d\epsilon_x}{(\mathrm{G}+\mathrm{H})\sigma_x-\mathrm{H}\sigma_y}=\frac{d\epsilon_y}{(\mathrm{F}+\mathrm{H})\sigma_y-\mathrm{H}\sigma_x}=\frac{d\gamma_{xy}}{\mathrm{N}\tau_{xy}}, \quad \ldots \quad (1)$$

where the x-axis is the direction of rolling and the y-axis is transverse to this in the plane of the sheet. The applied stress is considered to be in the plane of the sheet, with components σ_x, σ_y, and τ_{xy}. F, G, H, and

N are parameters specifying the current state of anisotropy, which will normally vary during a process of plastic distortion. Similar linear relations have been independently derived, though from different physical assumptions, by Jackson, Smith and Lankford (1949) and by Dorn (1949). Dorn's relations contain five parameters, and cannot be derived from a plastic potential (see equation (10) below).

(i) *Strain-ratios in a tension test.*—For a uniaxial tension σ applied at an angle α to the direction of rolling, the (x, y) components of stress are

$$\sigma_x=\sigma\cos^2\alpha, \quad \sigma_y=\sigma\sin^2\alpha, \quad \tau_{xy}=\sigma\sin\alpha\cos\alpha. \quad . \quad . \quad (2)$$

The strain-increment in the width direction of the tension specimen, with current width w, is

$$\frac{dw}{w}=d\epsilon_x\sin^2\alpha+d\epsilon_y\cos^2\alpha-2d\gamma_{xy}\sin\alpha\cos\alpha, \quad . \quad . \quad . \quad (3)$$

while the strain-increment through the thickness t is

$$\frac{dt}{t}=-(d\epsilon_x+d\epsilon_y) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

since there is no change of volume. Combining these four equations we obtain

$$\frac{dw/w}{dt/t}=\frac{\mathrm{H}+(2\mathrm{N}-\mathrm{F}-\mathrm{G}-4\mathrm{H})\sin^2\alpha\cos^2\alpha}{\mathrm{F}\sin^2\alpha+\mathrm{G}\cos^2\alpha}. \quad . \quad . \quad . \quad (5)$$

If the state of anisotropy is not altered by the deformation during the test, the ratios of the parameters remain constant and (5) can be integrated to give

$$\frac{\ln(w_0/w)}{\ln(t_0/t)}=\frac{\mathrm{H}+(2\mathrm{N}-\mathrm{F}-\mathrm{G}-4\mathrm{H})\sin^2\alpha\cos^2\alpha}{\mathrm{F}\sin^2\alpha+\mathrm{G}\cos^2\alpha}=r(\alpha), \text{ say}, \quad . \quad (6)$$

where w_0 and t_0 are the initial width and thickness. The ratio $r(\alpha)$ between the logarithmic strains in the width and thickness directions then has a constant value throughout any one test. In a completely isotropic material $\mathrm{F}=\mathrm{G}=\mathrm{H}=\mathrm{N}/3$ and $r(\alpha)\equiv 1$. If the sheet is isotropic only in its plane, $\mathrm{N}=\mathrm{F}+2\mathrm{H}=\mathrm{G}+2\mathrm{H}$, $r(\alpha)\equiv \mathrm{H}/\mathrm{F}$. In a non-isotropic material the ratios H/G and H/F can be found from tension tests along and perpendicular to the direction of rolling. Thus :

$$\frac{\mathrm{H}}{\mathrm{G}}=r(0), \qquad \frac{\mathrm{H}}{\mathrm{F}}=r(\pi/2). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

A third ratio, N/H say, can be obtained from a test in some other direction. The theoretical relation (6) is then completely known, and its validity can be checked by comparing it with the experimental results over the whole range of orientations.

Some general properties of (6) are worth noting. Stationary values of $r(\alpha)$ occur when $\alpha=0°$, $90°$, and also for a certain intermediate value β except when $(2\mathrm{N}-\mathrm{F}-\mathrm{G}-4\mathrm{H})/\mathrm{H}$ lies between $(\mathrm{F}-\mathrm{G})/\mathrm{G}$ and $(\mathrm{G}-\mathrm{F})/\mathrm{F}$.

The intermediate stationary value is a maximum when $(2N-F-G-4H)/H$ is positive, and a minimum when it is negative. In particular, when the properties are the same in both the rolling and transverse directions, so that $F=G$, the $r(\alpha)$ relation is symmetrical about $\beta=45°$; its value at the latter is a maximum if $N>F+2H$ and a minimum if $N<F+2H$.

In an anisotropic specimen a tensile stress generally produces a shear strain as well as an extension; that is, a principal strain axis does not usually coincide with the direction of the applied stress. It is easily shown from the above equations that the shear-strain increment, referred to axes along and perpendicular to the applied stress, is equal to

$$\frac{dw}{w} \times \frac{[(N-F-2H)\sin^2\alpha-(N-G-2H)\cos^2\alpha]\sin\alpha\cos\alpha}{[H+(2N-F-G-4H)\sin^2\alpha\cos^2\alpha]}. \quad . \quad (8)$$

The presence of a shear strain has been demonstrated experimentally by Hazlett, Robinson and Dorn (1949). In the present tests the angle of shear, as computed from (8) with the deduced values of the parameters, was never more than about 5°, even at the largest extensions. The corresponding eccentricity of loading was negligible in view of other experimental errors and the unavoidable scatter in the results. It can be seen from (8) that the axes of stress and strain-increment coincide when $\alpha=0°$, 90°, or an intermediate value γ satisfying

$$\tan^2\gamma=\frac{N-G-2H}{N-F-2H}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

γ coincides with the stationary point β of the $r(\alpha)$ relation only when $F=G$.

(ii) *Angles of necking.*—A *sufficiently thin* flat strip necks at an oblique angle when pulled in tension. In a perfectly isotropic specimen the angle is about 55° to the axis. According to theory (Hill 1948), the line of the neck develops in the direction of zero rate of extension in the uniformly deforming strip. In a non-isotropic specimen the two directions of zero extension-rate make different angles with the axis, except when $\alpha=0°$, 90°, or γ. If, then, the neck does in fact coincide with a direction of zero extension, a further check on the preceding theory is provided by measurements of the necking angles. For example, for a test in the rolling direction the theoretical inclination of the neck to the axis is $\tan^{-1}\sqrt{(1+G/H)}$.

Unfortunately the fractures in the copper and brass specimens used in the present experiments were too ragged for sufficiently precise measurements of the necking-angles. It appeared, however, that the necks usually formed at greater angles to the specimen axis than would be expected theoretically. This discrepancy is probably due to a slight necking in the width direction, which occurred towards the end of the considerable range of uniform strain and extended over a large part of the gauge length. The theory seems to give closer agreement in a metal where the range of uniform strain is comparatively small (Hill 1948).

(iii) *Earing positions.*—The state of stress and strain in a circular blank which is being deep-drawn into a cup can, in principle, be calculated from equations (1) along with some assumed law of work-hardening under combined stress. The positions and gradual development of the ears would be included in the results of the analysis. However, although the calculation can be made for an isotropic blank (Hill 1949) the analysis becomes extremely involved when the blank is anisotropic. Consequently, some simpler method must be adopted for predicting the earing positions.

In view of the approximate symmetry of the deformation about the tip of an ear or the bottom of a hollow, it seems a reasonable assumption (Hill 1948) that the ears and hollows develop at the points on the rim where the principal axes of the stress and the strain-increment coincide. Since the stress in the rim is a circumferential compression, it follows from the previous theory that the ears and hollows should begin to form at points on the rim whose angular positions are $0°$, $90°$ and $90°-\gamma$, relative to the direction of rolling. There are therefore four ears and four hollows (at most), and the equations (1) must be modified to represent the anisotropy of a material giving more than four ears (see below). Now, when the yield criterion and stress–strain equations are connected on the plastic potential basis, Hill has shown that the initial yield stress has stationary values in the directions for which the axes of stress and strain coincide. It is natural to assume, therefore, that the ears form in the positions where the (circumferential)yield stress is initially a minimum (numerically) and the hollows where it is a maximum. Then, if N is greater than both F+2H and G+2H the ears are in the $90°-\gamma$ positions, and if N is less than both F+2H and G+2H the ears are in the $0°$ and $90°$ positions. If N is intermediate to F+2H and G+2H only two ears are formed; these are in the $0°$ positions if F>G and in the $90°$ positions if F<G.

In particular, if F=G, there are four ears in the $45°$ positions if N>F+2H, and in the $0°$ and $90°$ positions if N<F+2H. These are also the positions where the strain-ratio r is a maximum in the special case F=G. This, indeed, is what would be expected if the circumferential strain in the blank did not depend on the angular coordinate, since the ears correspond to the parts of the blank which suffer most radial extension. When applied generally, this consideration leads to the hypothesis that the ears develop at positions where the strain-ratio r is a maximum for a uniaxial stress in the *circumferential* direction. However, this alternative hypothesis is unlikely to be correct since in actuality the circumferential strain varies considerably around the rim of a non-isotropic blank. In fact, the experimental evidence reviewed later is more in accord with the assumption that the ears form at positions where the strain-ratio r is a maximum for a uniaxial stress in the *radial* direction. This third hypothesis was suggested on empirical grounds by Baldwin Howald and Ross, but it is difficult to see a reasonable theoretical basis for it.

It is worth remarking that the absence of ears in a cupping test does not prove that the sheet is completely isotropic, but only isotropic in its

plane. According to the above theory this implies $N=F+2H=G+2H$, but H is not necessarily equal to F and G. To prove complete isotropy it is essential to perform, in addition, a tensile test in the plane of the sheet to see whether the strain-ratio is unity or not.

With suitable rolling and heat treatment some metals can be made to give six ears in a cupping test, *e. g.* brass, or even eight ears, *e. g.* certain aluminium alloys (Chevigny 1949). To represent these more complex states of anistropy it is necessary to generalize the stress–strain equations (1). This may be achieved (Hill 1950) by writing

$$\frac{d\epsilon_x}{\partial f/\partial\sigma_x}=\frac{d\epsilon_y}{\partial f/\partial\sigma_y}=\frac{d\gamma_{xy}}{\partial f/\partial\tau_{xy}}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

where f, the plastic potential, is taken to be a polynomial of degree n in the components of stress*. f is homogeneous if the strain-ratios do not depend on the absolute magnitude of the stress. The corresponding number of ears is $2n$ (or less, in special circumstances). The plastic potential for equations (1) is the quadratic

$$(G+H)\sigma_x^2-2H\sigma_x\sigma_y+(H+F)\sigma_y^2+2N\tau_{xy}^2. \quad . \quad . \quad . \quad . \quad (11)$$

To represent a material giving six ears, f must be a cubic (at least) and for a material giving eight ears f must be a quartic (at least).

§ 3. Materials and Treatments.

The materials used were copper and brass. In pure copper there are two common recrystallization textures, viz. (100) [001] giving four ears at 0° and 90°, and (113) $[\bar{2}1\bar{1}]$ giving four ears at 45°. Only the latter was studied here since the former has been thoroughly investigated by Baldwin, Howald and Ross. In 70/30 brass the usual earing positions are at 45° or at 0° and 60°. Both these were examined.

The copper was phosphorus deoxidized, with an average grain size of 0·06 mm. It was received as a hot-rolled and pickled strip of width 7 in. and thickness $\frac{3}{8}$ in. The strip was machined to 0·210 in. to remove surface pitting, and rolled to a nominal thickness of 0·030 in. (85 per cent reduction). It was then annealed at 650° C. for half an hour.

The brass was received as a hot-rolled strip of width 7 in. and thickness $\frac{3}{8}$ in. To obtain a texture giving four ears the strip was rolled 50 per cent, annealed at 425° C., rolled 84 per cent, and annealed at 650° C. The final nominal thickness was 0·030 in. To obtain a texture giving six ears the strip was rolled 50 per cent, annealed at 650° C., rolled 84 per cent, and annealed at 650° C. The final thickness was again 0·030 in. Burghoff and Bohlen (1942) have determined the (111) pole figures for both treatments.

* In the differentiation τ_{xy} and τ_{yx} are treated as distinct, the terms containing the shear–stress component being regarded as split into two equal parts. Only even powers of the shear–stress are allowable since the axes of anisotropy are the axes of reference.

In the cupping tests 4 in. diameter blanks were drawn with a punch of diameter 2 in. over a die of radius 0·187 in. The nominal blank-holder clearance was 0·003 in. The lubricant was a mixture of tallow and graphite. Cups drawn from the copper strip had four large ears at 45° to the rolling direction. The four-eared brass cups had large ears at approximately 50° to the direction of rolling, and the six-eared brass cups had small ears at 0° and 58° (±2°).

The tensile test pieces were produced in a blanking press and milled to size, the edges being polished on emery paper. Specimens (in duplicate) were cut at 0°, 22½°, 45°, 67½° and 90°. The gauge length was 2 in. and the width 0·5 in. The tests were carried out in a Hounsfield tensometer. Extensions in the elastic and initial plastic range were measured by a Lindley extensometer. Width and thickness measurements (average of three along the gauge length) were taken with a hand-micrometer reading to 0·0001 in. The readings were continued up to the stage at which necking began.

§ 4. Experimental Results.

With all the materials tested it was found that the strain-ratio r did not vary with strain, within experimental scatter. The constancy of r was maintained until necking intervened; this occurred at extensions of the order of 30–40 per cent. The inference to be drawn is that the initial anisotropy in the rolled strip was not appreciably altered by the uniform tensile strains.

Values of $r(\alpha)$, derived from the slopes of the best straight lines through the experimental points (means of at least two tests), are shown in fig. 1. Variations in the material, errors of measurement, and scatter in the results, combine to produce an uncertainty of order ±0·1 in the derived values of r. Ratios of F, G, H, N were obtained by fitting (6) to the results for the copper and brass giving four ears. There is some uncertainty in the values of the parameters, amounting perhaps to ±0·1, owing to the arbitrariness in choosing the best fit.

The curve in fig. 1 (*a*) for the copper corresponds to F=G=H=N/4. The theory in § 2 (iii) predicts four ears at 45° when F=G and N>F+2H, in agreement with experiment. Comparison may be made with the data of Baldwin, Howald and Ross, for copper giving 45° ears (see the lower curve in fig. 18 of their paper). The parameters for their material, which had received a different treatment, are F=G=H=N/5.

The parameters for the brass giving four ears, corresponding to the curve in fig. 1 (*b*), are

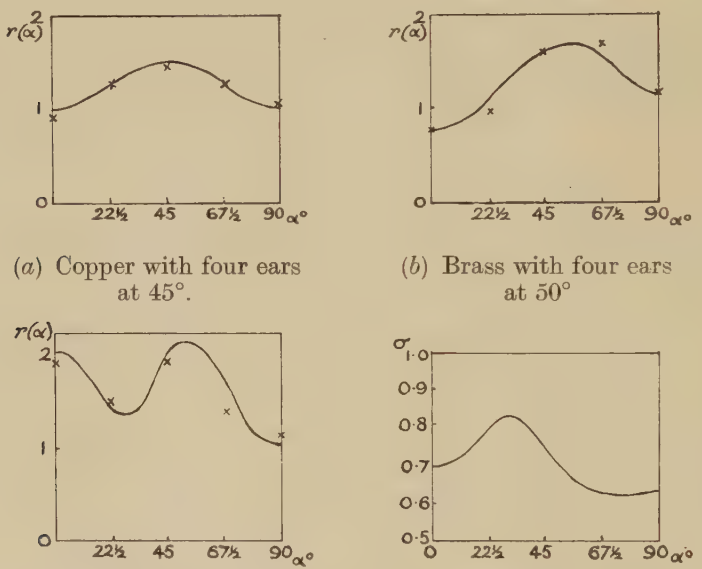
$$F : G : H : N = 1 : 1{\cdot}6 : 1{\cdot}2 : 5{\cdot}5.$$

The angle γ in (9) then has the value 40°, approximately, so that $90° - \gamma = 50°$. Since N is greater than both F+2H and G+2H, the theory predicts four ears at 50°. This is what was observed. It is, of course, hardly possible to be precise about the actual angular position of the tip

of an ear, especially since there are always slight differences between each of the four ears. In view of the uncertainty in the derived values of the parameters, a discrepancy of less than 5° between theory and experiment would probably not be significant.

As a further check on the validity of equation (1), a compression test normal to the sheet was carried out on a cylindrical brass disc of diameter 0·25 in. and thickness 0·029 in. The section became elliptical, with the major axis in the direction of rolling. After a 50 per cent compression the axes were 0·38 in. and 0·325 in. (approximately, since the edge became ragged)*. The corresponding ratio of the logarithmic strains is 1·6.

Fig. 1.



(*a*) Copper with four ears at 45°.

(*b*) Brass with four ears at 50°

(*c*) Brass with six ears at 0° and 58°.

(*d*) Theoretical variation of yield stress with orientation in brass with six ears.

Regarding the test as equivalent to the application of equal biaxial tensions in the plane of the sheet, equation (1) gives $d\epsilon_x/d\epsilon_y = G/F = 1{\cdot}6$, in perfect agreement.

In fig. 1 (*c*), for the brass giving six ears at 0° and 58°, the curve corresponds to the relation

$$r(\alpha) = \frac{6\cos^6\alpha + 6\sin^6\alpha + 38\sin^4\alpha\cos^2\alpha - 10\sin^2\alpha\cos^4\alpha}{3\cos^4\alpha + 6\sin^4\alpha + \sin^2\alpha\cos^2\alpha}.$$

This was obtained by assuming a homogeneous cubic polynomial for the plastic potential defined in equation (10), and choosing the ratios of the six numerical coefficients so that the following five conditions were

* Calcium oleate was the lubricant (coefficient of friction ~0·015). With these specimen dimensions the ellipticity would not be significantly influenced by friction.

satisfied : (i) $r=2$ in the rolling direction ; (ii) $r=2$ in the 45° direction ; (iii) $r=1$ in the transverse direction ; (iv) the axes of stress and strain to coincide in a tension test at 30° to the direction of rolling ; (v) the ratio of the strain in the rolling direction to the transverse strain to be 1·5 in a compression test normal to the sheet. The latter is regarded as equivalent to equal biaxial tension in the plane of the sheet. The value 1·5 is near to that observed in a compression test on a cylindrical disc of diameter 0·25 in. and height 0·03 in., lubricated with calcium oleate. After a 50 per cent compression the axes of the ellipse were 0·375 in. and 0·325 in., approximately.

The plastic potential so determined has the form

$$f\equiv 3\sigma_x^3-6\sigma_x^2\sigma_y-6\sigma_x\sigma_y^2+4\sigma_y^3+(4\sigma_x+21\sigma_y)\tau_{xy}^2. \quad . \quad . \quad . \quad (12)$$

According to the hypothesis of §2 (iii) the fourth condition ensures that either an ear or a hollow occupies the 60° position. The other possible positions of ears and hollows are calculated to be 0°, 90°, and $\cot^{-1}\sqrt{11}=17°$ approximately. It will be seen that the calculated curve passes close to the two experimental points not deliberately fitted, viz. at $22\frac{1}{2}°$ and $67\frac{1}{2}°$. This method of comparison was preferred to the obvious one of fitting the curve to all five points, and then comparing predicted and observed earing positions, because there is considerable latitude in fitting a cubic curve to so few points when they are uncertain to $\pm 0{\cdot}1$.

It remains to distinguish between the calculated positions of the ears and hollows. If the yield criterion appropriate to (12) is $f=$constant (depending on the strain-history), the ears are predicted at 0° and 60° if it is supposed that they form at points where the circumferential yield stress is a (relative) *maximum*. This is opposite to what was found above for copper and brass sheet giving four ears, where the ears form at points where the circumferential yield stress is, theoretically, an (absolute) minimum. We are unable to suggest an explanation of the difference ; it may be an accidental product of the theory, depending in a complex way on the entire distribution of the stresses and strains in the initial stages of the cupping test. The calculated variation of the yield stress with the orientation of the tension specimen is shown in fig. 1 (*d*).

It is interesting to notice that the earing positions are close to the angles for which the strain-ratio r has maximum values for a *radial* tension. This is the hypothesis of Baldwin, Howald and Ross, referred to previously. The present experiments do not indicate whether this is the real factor determining earing, or whether it is merely that the maximum values of r happen to fall rather close to the actual earing positions on account of some more fundamental property (which may, indeed, be peculiar to brass and copper).

The variation of the strain-ratio r with orientation has been measured for an aluminium alloy (61S–T6) by Hazlett, Robinson and Dorn (1949). The composition was 0·25 per cent Cu, 0·6 per cent Si, 1·0 per cent Mg, and 0·25 per cent Cr. The alloy was solution heat-treated, quenched, and aged. Values of the parameters giving best fit of equation (6) to

their data are $F : G : H : N = 1 : 1{\cdot}15 : 0{\cdot}65 : 2{\cdot}25$. Since N is only slightly different from F+2H or G+2H the shear strains (8) should be negligible, and this was found to be so. The earing positions were not measured. The same authors found that the strain-ratios for a magnesium alloy with only slight anisotropy varied during a tension test. Measurements of strain-ratios in an aluminium alloy, for tests both in and out of the plane of the plate, have been reported by Klingler and Sachs (1948).

Strain-ratios for low carbon steel have been measured by Lankford, Snyder and Bauscher (1950), who make the interesting suggestion that plastic anisotropy may sometimes be advantageous to good performance in an *unsymmetrical* pressing operation.

§5. Stress–strain Curves.

For the correlation of stress–strain curves, obtained in tension tests at various angles to the direction of rolling, it is necessary to relate some measure of the amount of work-hardening to the previous strain-history. Assuming that the amount of hardening can be specified by a scalar quantity (which appears probable when the state of anisotropy does not vary), the natural measure to take is the value of the expression f, representing the yield criterion and plastic potential. Jackson, Smith and Lankford (1949) claim to have found evidence that the degree of hardening in anisotropic sheet steel is a function only of the amount of plastic work per unit volume (a hypothesis which is a reasonable approximation for some isotropic metals.) The present tests do not support this hypothesis in the cases of copper and brass.

Thus, the stress–strain curves at 0°, $22\frac{1}{2}$°, 45°, $67\frac{1}{2}$°, and 90° for the brass giving four ears did not vary from a mean by more than 0·8 tn/in.^{2} at any strain up to 50 per cent extension. In view of the variation between nominally similar specimens, there was scarcely any significant difference between the stress–strain curves in the various directions. Hence, if the hardening depended only on the amount of plastic work, the tensile yield stress should not vary significantly with direction in a given mechanical state. This clearly follows from the hypothesis that the same mechanical state can be produced by performing the same amount of work in tension tests in different directions; since the stress–strain curves coincide, the final yield stress is the same in all directions. This conclusion, however, is not consistent with the assumption that the yield criterion is f=constant, since this implies variations of the yield stress with direction. For example, according to (11), the ratio of the yield stresses in the 0° and 90° directions should be $\sqrt{\{(F+H)/(G+H)\}}$ or about 0·9; the ratio of the yield stresses in the 40° and 90° directions is 0·8, approximately.

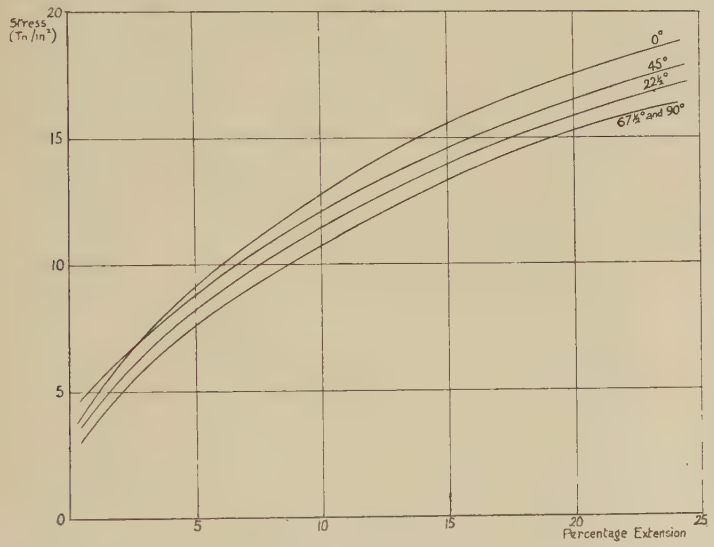
A similar conclusion applies also to the brass earing at 0° and 60°, where the stress–strain curves showed even less variation and the theoretical yield stresses more.

The stress–strain curves for the copper, however, showed significant differences (fig. 2), especially between the 0° and 90° curves. Since the theoretical yield stresses are roughly equal at 0° and 90°, the same conclusion is again reached, namely that the hardening (as measured by the value of f) is not a unique function of the total plastic work per unit volume.

Acknowledgment.

We are grateful to the Director of the British Iron and Steel Research Association for permission to publish this work.

Fig. 2.



Stress–strain curves for copper at various inclinations to the rolling direction (experimental points omitted for clarity).

References.

Baldwin, W. M., Jr., Howald, T. S., and Ross, A. W., 1946, *Trans. Am. Inst. Min. Met. Eng.*, **166,** 86.

Burghoff, H. L., and Bohlen, E. C., 1942, *Trans. Am. Inst. Min. Met. Eng.*, **147,** 144.

Chevigny, R., 1949, *Revue de l'Aluminium,* March, p. 79 ; 1950, *Sheet Metal Industries*, **27,** 133.

Dorn, J. E., 1949, *Journ. App. Phys.*, **20,** 15.

Hazlett, T. H., Robinson, A. T., and Dorn, J. E., 1949, University of California, Engineering Research Project N6-ori-111, Task No. 4, May.

Hill, R., 1948, *Proc. Roy. Soc.* A, **193,** 281 ; 1949, British Iron and Steel Research Association, Report MW/E/48/49, Sept. ; 1950, *The Mathematical Theory of Plasticity*, Chap. XII, (Clarendon Press : Oxford).

Jackson, L. R., Smith, K. F., and Lankford, W. T., 1948, *Metals Technology, Tech. Pub.* 2440, August ; see also discussion of this paper in 1949, *Journal of Metals*, **1,** 324.

Klingler, L. J., and Sachs, G., 1948, *Journ. Aero. Sci.*, **15,** 599.

Lankford, W. T., Snyder, S. C., and Bauscher, J. A., to be published in *Trans. Am. Soc. Metals.*

LXI. *The Fluctuation and Fading of Radio Echoes from Meteor Trails.*

By J. S. Greenhow,
University of Manchester *.

[Received April 21, 1950.]

[Plates XXII. & XXIII.]

Abstract.

This paper is concerned with the amplitude fluctuations observed in the radio waves reflected from meteor trails. Subsequent to the initial diffraction phenomena the radio echoes exhibit fluctuations of period between 0·01 and 0·1 seconds, when observed on radio wave frequencies of 36 to 72 Mc./s. These fluctuations have been investigated simultaneously on two radio wave frequencies and it is concluded that the phenomenon is due to wind gradients in the high atmosphere which cause the trail of meteor ionization to break up. In this way several reflecting regions are formed which give an interference pattern moving over the observing station.

§ 1. Introduction.

Observations of the radio echoes from meteor trails indicate that the amplitude fluctuations during the life of the echo are complex in character. For example, during the Giacobinid shower of October 1946, Lovell, Banwell and Clegg (1947) investigated the life histories of several echoes and found that many showed rapid fluctuations in amplitude superimposed on the general decay of the echo. Similar effects have been found by McKinley and Millman (1949). These observations were made using either cine-photography or photography of intensity modulated displays, and it was evident that further progress in the study of this phenomenon could only be made by investigating the amplitude of the echo over much shorter time intervals. This paper describes the information which has now been obtained on the fluctuations using a technique which enables the amplitude of successive radio pulses to be measured with a time separation of 6·6 milliseconds.

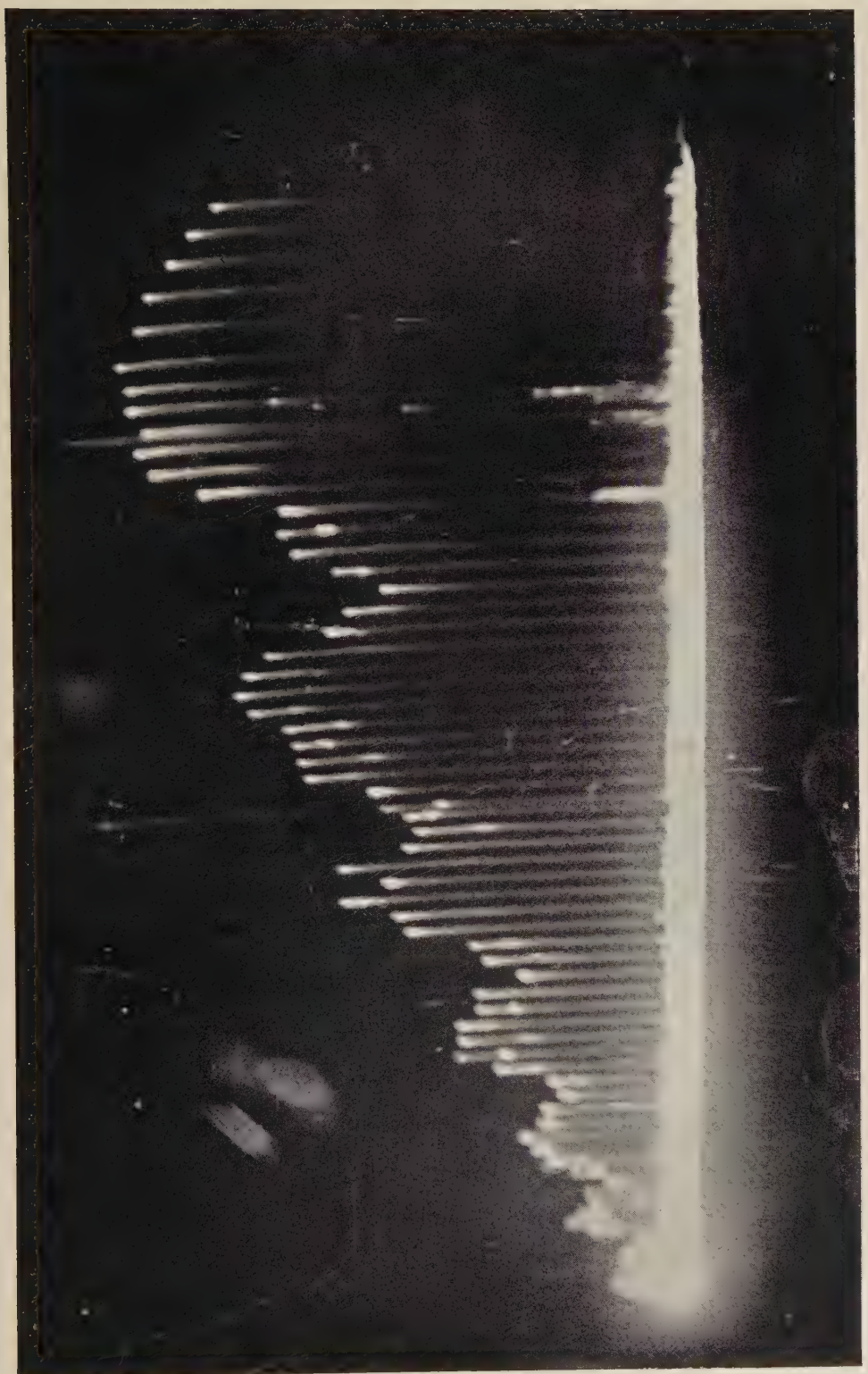
During the complete life history of a characteristic radio echo from a meteor trail it is possible to distinguish four types of fluctuation.

(*a*) *Diffraction Phenomena.*

Approximately 10 per cent of meteor echoes show a rapid initial rise in amplitude, followed by oscillations of increasing frequency. A typical echo of this type is shown in fig. 1 (Pl. XXII.). The oscillations, which are usually

* Communicated by Dr. A. C. B. Lovell.

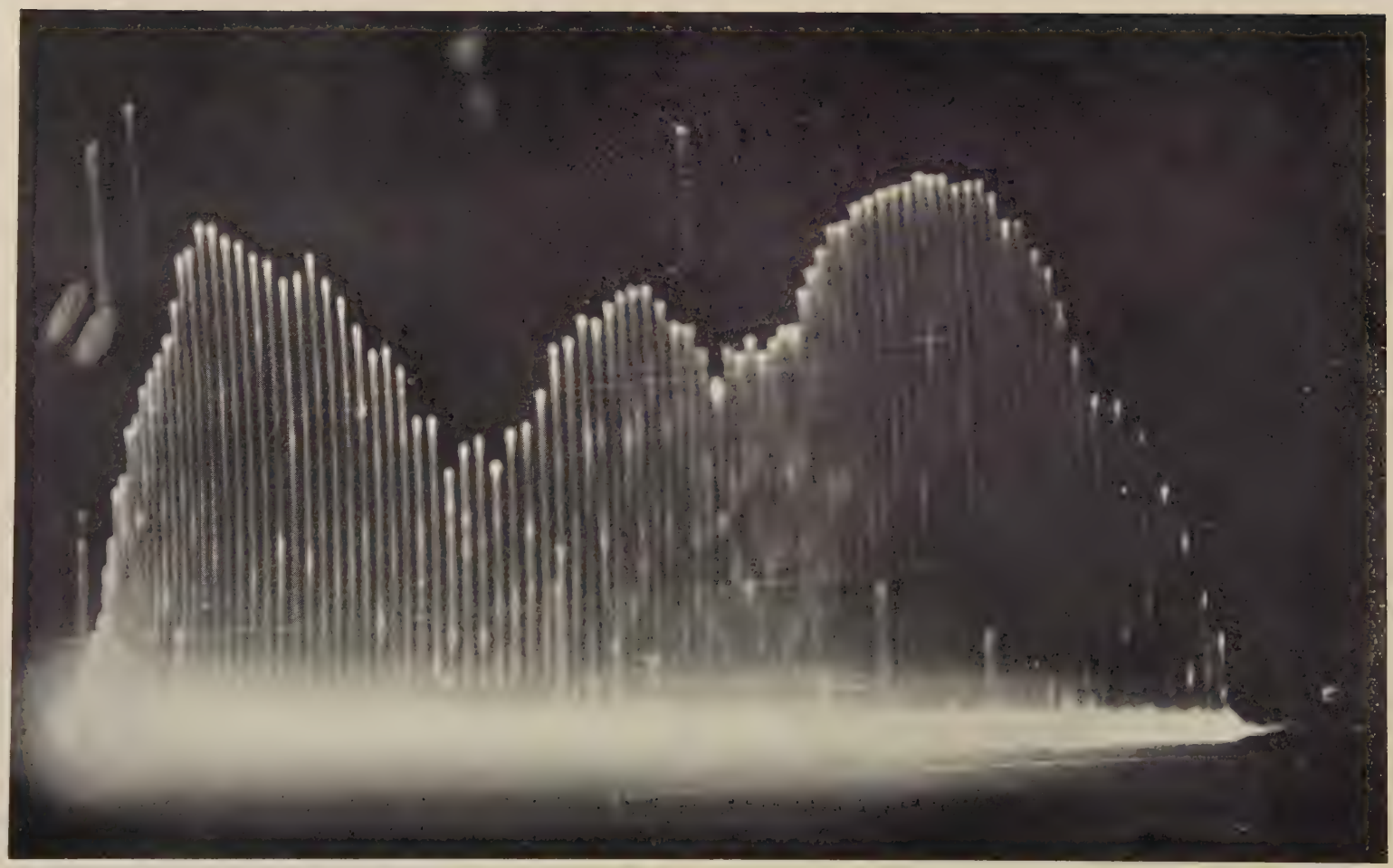
← Amplitude



← Time

Fig. 1. Typical diffraction pattern. Velocity of meteor: 340 km/sec.

FIG. 2



Example of short period flutter. Pulse repetition frequency: 750 c/s. Mean Period of flutter: 0·04 seconds.

complete with 0·5 seconds of the first recorded pulse correspond to the passage of the head of the meteor through successive Fresnel diffraction zones, and they can be used to measure the meteor's velocity (Ellyett and Davies 1948, Davies and Ellyett 1949).

(*b*) *Short Period Flutter.*

Although the majority of echoes have durations of less than 0·2 sec., some 25 per cent persist for much longer periods, and show violent fluctuations of the type illustrated in fig. 2 (Pl. XXIII.). These oscillations, which have periods of the order 0·01 to 0·1 seconds, are occasionally observed directly following the Fresnel fluctuations, but more usually occur with echoes where the diffraction phenomena are not observed.

(*c*) *Long Period Fluctuations.*

In a small percentage of very long duration echoes, a much slower amplitude variation, of the order of 1·0 seconds period, has been detected.

(*d*) *Noise Fluctuations.*

In addition to the foregoing types of amplitude oscillation all echoes invariably show small fluctuations between adjacent pulses, which are due to random noise in the receiver (Ellyett 1949).

The phenomena described in (*a*) and (*d*) are now well understood, and will not be discussed further in this paper. The short period flutter (*b*) and long period fluctuations (*c*), however, present a number of interesting physical problems, and have been investigated in detail.

§ 2. Technique.

Most of the results have been obtained by a dual wave technique, giving continuous pulse by pulse amplitude records on 72 and 36 Mc./s. simultaneously, for the whole duration of an echo. Additional information has been obtained (on 72 Mc./s. only), from earlier techniques, using automatic photography of successive pulse amplitudes from sections of the echo.

(*a*) *Dual wave technique.*

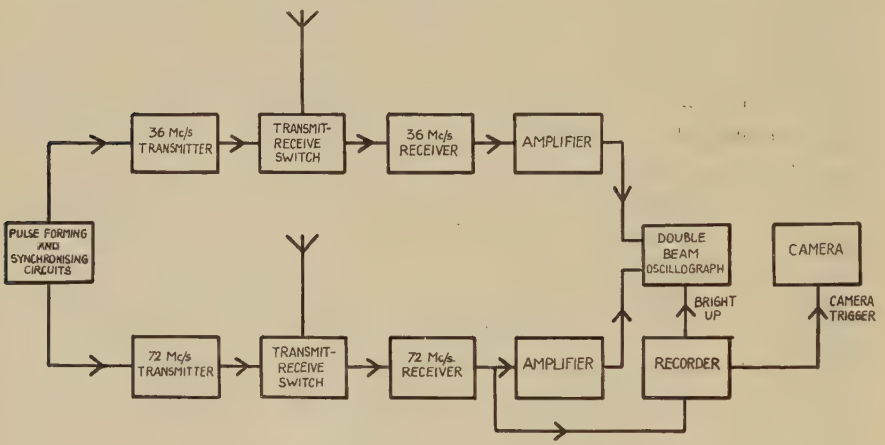
This technique was developed by Ellyett * and a block diagram of the equipment is shown in fig. 3. Two transmitting and receiving systems were used, operating at frequencies of 72 and 36 Mc./s. Upon receipt of the first pulse exceeding twice the signal to noise ratio reflected from a meteor trail, a recorder unit (Davies and Ellyett 1949), attached to the 72 Mc./s. receiver output, actuated the dual wave unit. Both receiver outputs were fed continuously via separate amplifying channels in this unit, to a double beam cathode ray tube. A camera with continuously moving film was switched on simultaneously, so that successive pulses

* Dr. C. D. Ellyett returned to New Zealand before the final completion of the dual wave technique.

were separated on the film, giving a pulse by pulse picture of the echo amplitude throughout its life. The camera motor required approximately 0·05 sec. to reach a speed sufficient to separate the individual pulses, so this period was lost at the beginning of each echo. Once the recorder had actuated the system a hold-on relay came into operation, maintaining the unit in action, until switched off manually on conclusion of the echo.

The two transmitters were locked together to the same pulse repetition frequency, and their peak powers were adjusted so that echoes appearing on both wavelengths simultaneously had approximately the same signal to noise ratios. The aerials were horizontal folded dipoles with reflectors, at heights of half a wavelength above the ground, and looking in a northerly direction. Common transmit-receive systems enabled the same aerials to be used for both transmitting and receiving. The pulse repetition frequency used in these experiments was 150 per sec.

Fig. 3.



Block diagram of the dual wave apparatus.

(*b*) *Single wave technique.*

An automatic recorder unit designed for observation of the Fresnel zone pattern has been described previously (Davies and Ellyett 1949). This unit photographs some 100 pulses on one photographic frame, and takes 1 sec., to reset. For long duration echoes detailed amplitude information can therefore be obtained from short sections of the echo separated by intervals of 1 second.

§ 3. The Short Period Flutter.

Photographic records of short period flutter have been obtained during most of the major showers of 1948 and 1949, although the dual wave technique was used only during the 1949 Geminid shower. Examination of the results shows that the individual fluctuations which comprise the flutter vary continuously in amplitude and frequency during the lifetime of an echo. The results have been analysed to find,

(*a*) the relationship between the echo duration and the mean time period of the flutter ;

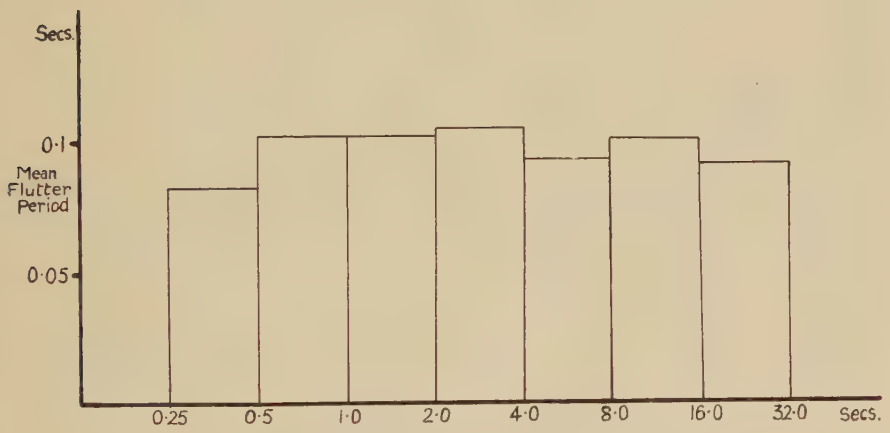
(*b*) the relationship between the mean flutter periods observed on frequencies of 36 Mc./s. and 72 Mc./s. ; and

(*c*) the relationship between the flutter period and the type of meteor.

(*a*) *Relation between Echo duration and Mean Period of Flutter.*

A histogram relating the mean period of the flutter to the echo duration, for echoes observed on a frequency of 36 Mc./s., is shown in fig. 4. It is clear that the mean period is practically the same for each duration group. The processes causing the phenomenon of high speed flutter must therefore be independent of the duration of the echo.

Fig. 4.



Echo duration. (36 Mc./s.).

Mean flutter period as a function of echo duration.

(*b*) *Relationships between the Mean Periods of Fluctuation on frequencies of* 38 *Mc./s. and* 72 *Mc./s.*

The relationship between the flutter periods on 36 Mc./s. and 72 Mc./s. was investigated with the aid of the dual wave apparatus.

Table I. gives the mean flutter periods for the two frequencies, and the ratios between these periods.

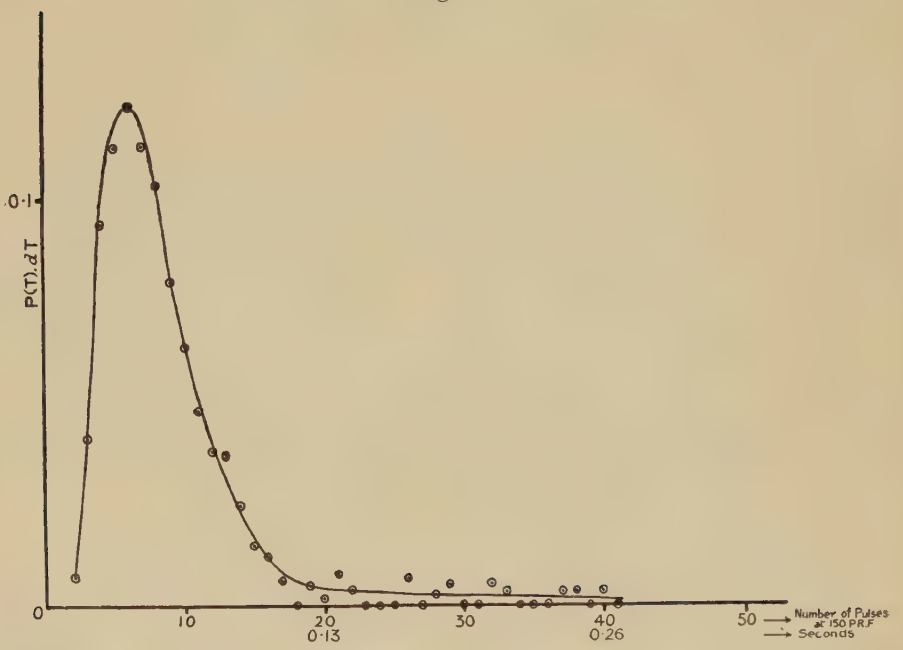
TABLE I.

	Mean flutter period (secs.)	Ratio of the mean flutter period at 36 Mc./s. to the mean flutter period at 72 Mc./s.
72 Mc./s. echoes	0·045	1·0
36 Mc./s. echoes while accompanied by a 72 Mc./s. echo	0·082	1·84
All 36 Mc./s. echoes	0·091	2·02

It is evident that the period of the oscillations on a frequency of 36 Mc./s. is approximately twice that of those observed on 72 Mc./s. In general the echo observed on a wavelength of 4 m. is of much shorter duration than the echo observed on 8 m. from the same trail (Lovell 1948) but there is no significant difference between the mean periods on 36 Mc./s. before and after the end of the corresponding 72 Mc./s. echo.

Since the flutter period is independent of the duration of the echo, the probability P(T) of a fluctuation of period T occurring is proportional to T . N(T), where N(T) is the total number of fluctuations of this period observed. The relationship between T . N(T) and T for the echoes photographed on the 72 Mc./s. apparatus is shown in fig. 5. This is compared with the corresponding curve for 36 Mc./s. echoes in fig. 6.

Fig. 5.



Flutter period (T).

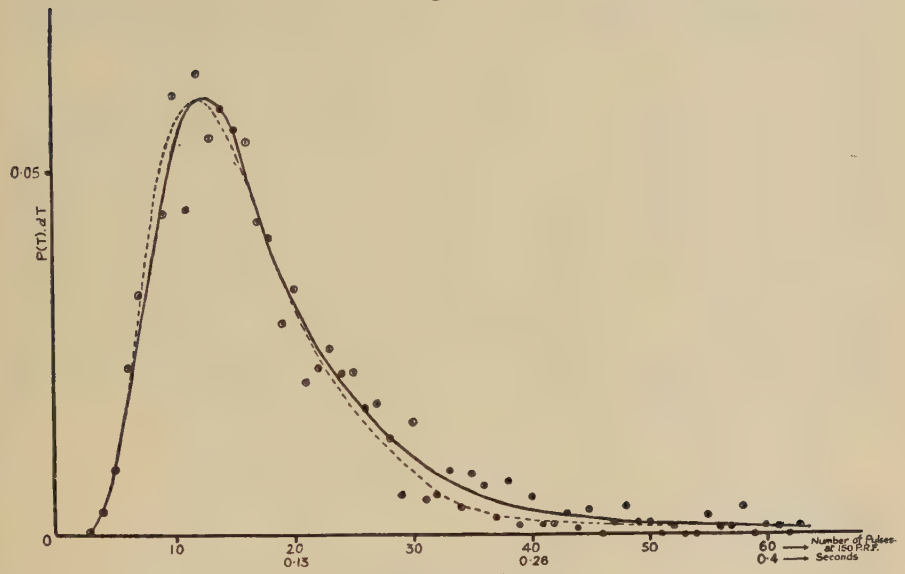
P(T)dT: Probability of a fluctuation of period between T and T+dT occurring at 72 Mc./s. (dT=0·066 secs).

If the mean flutter periods T_1 and T_2 observed on the 36 Mc./s. and 72 Mc./s. equipments are related to the wavelengths by a power law, the exponent n will be given by

$$n = \frac{\log_{10} T_1/T_2}{\cdot 301}.$$

This exponent was found for 36 echoes observed simultaneously on both wavelengths. Measurements were restricted to times when both echoes were present, and the distribution of values obtained for n is shown in

Fig. 6.

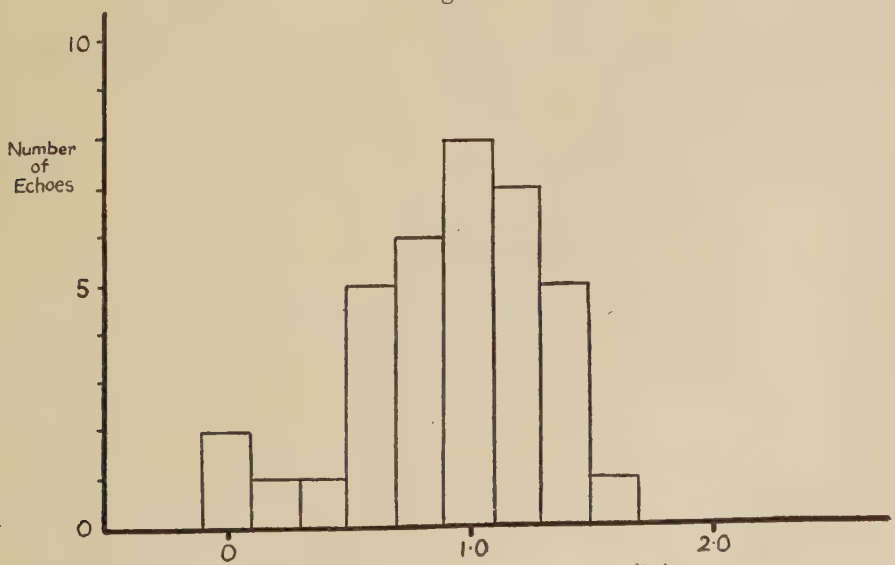


Flutter period (T).

(Continuous curve). P(T)dT : probability of a fluctuation of period between T and T+dT occurring at 36 Mc./s. (dT=0·066 secs.)

(Dotted curve). Corresponding curve for 72 Mc./s. scaled up by a factor of 2.

Fig. 7.



Exponent (n) calculated from $\frac{T_1}{T_2} = \left(\frac{\lambda_1}{\lambda_2}\right)^n$

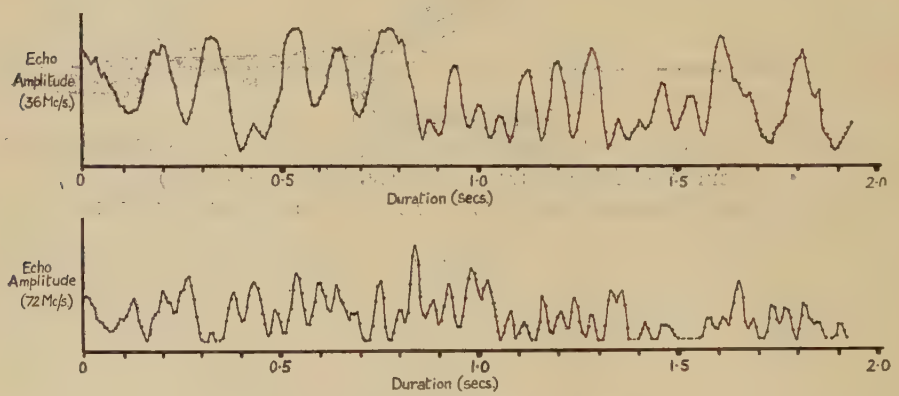
T_1 and T_2 are mean flutter periods at wavelengths of λ_1 and λ_2.

Distribution curve for echoes present simultaneously on wavelengths of 8·4m. (λ_1) and 4·2m. (λ_2).

fig. 7. The mean value was found to be 0·99. It therefore seems reasonable to conclude that the periods are proportional to the wavelength of the radiation employed.

Thirty-six echoes exceeding 0·5 sec., in duration at 72 Mc./s., whose amplitudes remain at a reasonable value for a sufficiently long period, have been analysed pulse by pulse. In general, the 72 Mc./s. fluctuations are of shorter periods than those at 36 Mc./s., and there are sections in most of the echoes where the ratio of the periods is exactly two to one. Fig. 8 shows an echo in which a high degree of correlation persists for an appreciable part of the echo life. In this example amplitude maxima at the higher frequency correspond approximately to the maxima and minima at 36 Mc./s. The echo lasted for 1·9 sec. at 72 Mc./s. and 5·8 secs. at 36 Mc./s., only the first 2·0 secs. of the 36 Mc./s. echoes are shown. The individual pulse amplitudes have been smoothed by a sliding group of two.

Fig. 8.



Time: 05.21 hrs. Dec. 13th 1949.

Pulse repetition frequency	150 c/s.
Duration of 72 Mc./s. echo	1·9 secs.
Duration of 36 Mc./s. echo	5·8 secs.

Individual pulse amplitudes smoothed by a sliding group of 2.

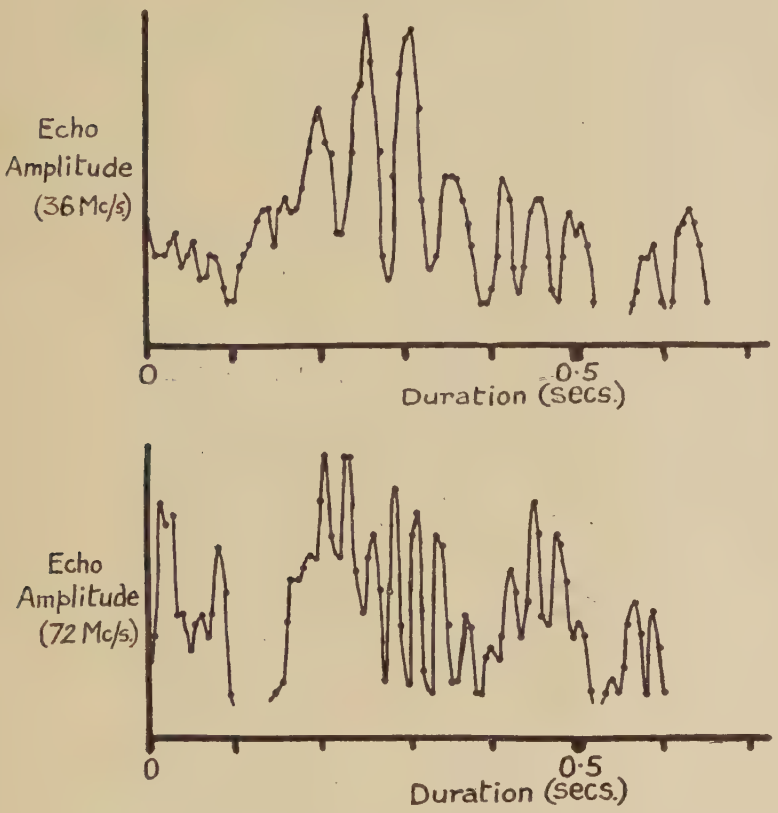
Fig. 9 is an echo of shorter duration showing higher speed fluctuations. The pulse amplitudes are unsmoothed. Fig. 10 is an example of an echo in which the fluctuations are less clearly defined. There is a two to one ratio in flutter periods for certain sections of the echo, but in other parts the fluctuations are confused and often of small amplitude. In general the fluctuations at 72 Mc./s. are faster than those at 36 Mc./s. The first 3·9 sec., of the echo are shown on both frequencies, although the 36 Mc./s. echo persists for a further 14 sec.

(*c*) *Relation between flutter period and types of meteor.*

In order to discover whether the flutter period is related to the type of meteor by which the trail was originally produced, or if it depends only upon the subsequent behaviour of the completed column of ionization, the

mean flutter periods for three major showers, and for sporadic meteors have been measured on 72 Mc./s. only, using the single wave technique. The results are given in Table II.

Fig. 9.



Time 22.24 hrs. Dec. 13th 1949.

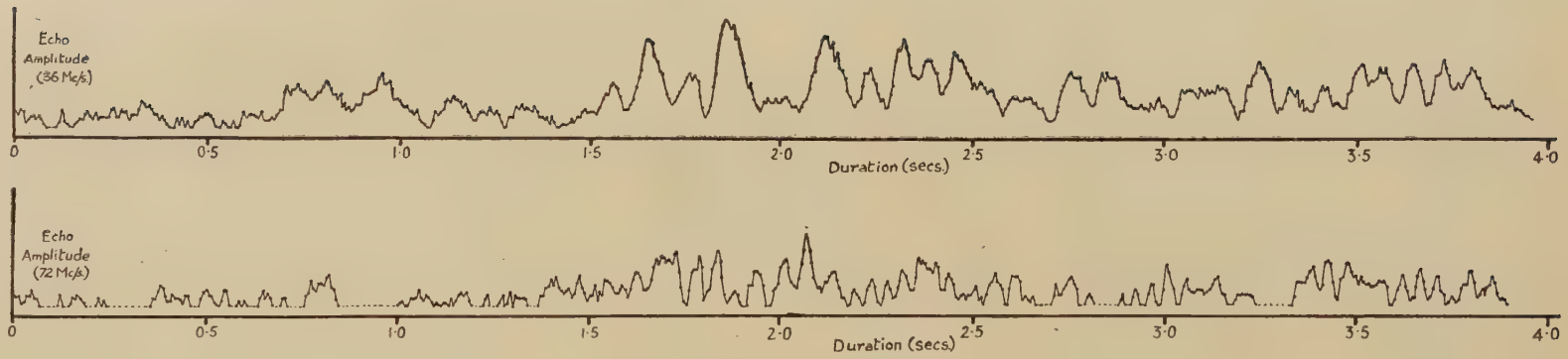
Pulse repetition frequency	150 c/s.
Duration of 72 Mc./s. echo	0·6 secs.
Duration of 36 Mc./s. echo	2·0 secs.

TABLE II.

Mean Period of Flutter (72 Mc./s.).

Shower	No. of observations of period	Mean period of flutter (sec.)
Perseids (Aug. 1948)	80	0·048
Geminids (Dec. 1948)	362	0·032
Quadrantids (Jan. 1949)	74	0·035
Perseids (Aug. 1949)	353	0·039
Sporadic meteors	41	0·033

Fig. 10.



Time : 0046 hrs Dec. 13th 1949.
Pulse repetition frequency 150 c/s.
Duration of 72 Mc./s. echo 3·9 secs.
Duration of 36 Mc./s. echo 17·1 secs.

It is evident that the mean period is independent of the type of meteor which produces the initial column of ionization.

§ 4. Long Period Oscillations.

Some experimental evidence for the existence of a longer period amplitude fluctuation has been found on 36 Mc./s. only. Out of a total of 30 echoes obtained by Ellyett * during the 1949 Perseid meteor shower, two, with durations of 81·0 sec. and 50·0 sec., show long period fluctuations emerging after about 30·0 sec., the period of the fluctuations gradually increases throughout the remainder of the echo life, and the mean period is of the order of 1·0 sec.

Similar behaviour has not yet been found using the dual wave equipment, owing to the shorter durations of the echoes on 72 Mc./s., but evidence of waves in one 7·0 sec. echo on 72 Mc./s. has been given by Lovell, Banwell and Clegg (1947), using cine-photography. Further information on the behaviour of the 72 Mc./s. echoes was obtained from a total of 90 cine records of the summer daytime and Perseid meteors of 1947, which show 10 cases of possible wave occurrence. A detailed analysis of the periods of the oscillations cannot be made from these records, owing to the possibility of a beating effect with the camera shutter mechanism.

As there are only two positive cases of the occurrence of long period amplitude fluctuations at 36 Mc./s., and none at 72 Mc./s., it is not yet possible to decide whether this is a separate phenomenon, or merely the tail of the distribution curve shown in fig. 4.

§ 5. Theoretical Explanations of the High Speed Amplitude Fluctuations.

Experimental evidence from several sources (for summary see Ellyett 1950) shows that there are both steady winds, with velocities of the order of 50–100 m./sec., and winds of a turbulent character, present in the region 70–120 km. above the earth.

Several suggestions have been made that the phenomenon of high speed flutter might be explained by the presence of such wind gradients in the upper atmosphere. (Herlofson 1948, Ellyett 1949, McKinley and Millman 1949). The results obtained in this paper will now be considered in terms of such wind motions. When a meteor trail is first formed, the ionization along the track is approximate linear, and it is known that the reflection of radio waves at 72 Mc./s. from the trail is specular (Lovell, Banwell and Clegg 1947, Hey and Stewart 1949). If the trail is subsequently deformed by winds, then other sections could present right angle reflecting points to the observer. An interference pattern due to the interaction of waves

* Unpublished.

reflected from various sections of the trail will then be formed on the surface of the earth. If the reflecting centres move relatively to each other, this interference pattern will move over the ground, producing a series of maxima and minima in the echo amplitude at any point.

If two sections of the trail predominate, and reflect approximately equal amounts of energy in the direction of the receiver, large amplitude fluctuations will be produced. The frequency of the fluctuations will be proportional to the frequency of the radio waves. The maximum possible flutter frequency will occur when the reflecting points are moving in opposite directions along the line of sight of the observer, with the maximum wind velocities present in the upper atmosphere. Any other flutter frequency less than this maximum is possible.

If V is the component of the velocity of the sources relative to each other in the direction of the observer, λ the wavelength, and T the period of the flutter, then

$$V \sim \frac{\lambda}{4T}.$$

V is independent of the wavelength, so that the flutter period T should be proportional to λ.

The distribution curve for all flutter periods at 72 Mc./s., scaled up by a factor of two is shown together with the corresponding distribution for 36 Mc./s. in fig. 2. (Pl. XXIII.). There is a reasonable correlation between the two curves. The most probable value of V is approximately 25 metres/sec., while the highest speed fluctuations correspond to a value of V of the order of 100 metres/sec. The slowest fluctuations observed have periods of approximately 0·5 sec. at 36 Mc./s. and 0·25 sec. at 72 Mc./s.

If more than two reflecting centres, of differing amplitudes, and moving with different velocities, are present, the fluctuations will become more complex, and correlations on widely different frequencies could not be expected. The mean flutter period, however, would still be proportional to the wavelength.

Additional information supporting the wind theory has been obtained by the visual observation of an auxiliary cathode ray tube. In several cases an echo was observed at 36 Mc./s. a few seconds before the appearance of the 72 Mc./s. echo. This suggests strongly that the track is undergoing distortion, and that a favourable electron distribution in the trail has temporarily caused the 72 Mc./s. echo to appear.

Acknowledgments.

The author wishes to thank the Director and Staff of the Manchester University Jodrell Bank Experimental Station, where these experiments were carried out; especially Dr. C. D. Ellyett, who analysed most of the single wave records and the cine records referred to in § 4, and Mr. R. L. Closs, who assisted materially in the taking of the dual wave observations. Valuable advice was given by Dr. A. C. B. Lovell, Director of the Station, and by Dr. J. A. Clegg.

Much of the work described in this paper has been made possible by the grant from the Department of Scientific and Industrial Research for the development of the experimental work at Jodrell Bank. The author is also indebted to the D.S.I.R. for the financial aid given during the course of this work.

References.

Davies, J. G., and Ellyett, C. D., 1949, *Phil. Mag.*, **40,** 614.
Ellyett, C. D., 1949, *Ph.D. Thesis*, University of Manchester; 1950, *Phil. Mag.*, **41,** 694.
Ellyett, C. D., and Davies, J. G., 1948, *Nature, Lond.*, **161,** 596.
Herlofson, N., 1948, *Rep. Prog. Phys.*, **11,** 444 (London: Physical Society).
Hey, J. S., and Stewart, G. S., 1947, *Proc. Phys. Soc.*, **59,** 858.
Lovell, A. C. B., 1948, *Rep. Prog. Phys.*, **11,** 415 (London: Physical Society).
Lovell, A. C. B., Banwell, C. J., and Clegg, J. A., 1947, *Mon. Not. R. Astr. Soc.*, **107,** 164.
McKinley, D. W. R., and Millman, P. M., 1949, *Proc. Inst. Radio Engrs.*, **37,** 364.

LXII. *The Influence of High Altitude Winds on Meteor Trail Ionization.*

By C. D. Ellyett, Ph.D. *,
University of Manchester †.

[Received April 21, 1950.]

Abstract.

The paper contains a discussion of the possible effect of high altitude winds on the ionized columns created by meteors. It is shown that the various types of range drifts observed in the long duration radio echoes from meteor trails can be explained in terms of the influence of winds on the ionized column. A comparison of the characteristics of the fading of the radio waves returned from the ionosphere, with the characteristics of the fluctuations of the radio echoes from meteor trails, leads to the conclusion that some of the ionospheric phenomena may be due to meteor ionization in the E region.

§ 1. Introduction.

There is now much evidence from the study of the radio wave reflections from meteor trails to show that the trail undergoes complex disturbances after its formation, and several authors have suggested that winds at altitudes of 80–120 km. are an important disturbing factor (Herlofson 1947, Hey and Stewart 1947, McKinley and Millman 1949). The purpose of the present paper is to examine the available evidence and to estimate the extent to which winds can account for the observed phenomena. The principal phenomena involved are those of range drifts and amplitude fluctuations found in the long duration echoes from meteor trails. The paper includes a discussion of the correlation of these effects with recent work on the fading of radio waves from the ionosphere (Ratcliffe 1948).

§ 2. Wind Processes at 70–100 km. Height.

Evidence from the movement of polar auroræ, of noctilucent clouds, and of meteor dust trails seen by reflected sunlight, gives conclusive proof of wind motion at heights of 70–100 km. (Störmer 1939). The mean drift velocity is of the order of 50–100 m./sec. Similar values for the wind velocities have been obtained recently by Mitra (1949) from direct ionospheric radio measurements.

In addition to steady winds, turbulent motion also occurs. Olivier (1947), analysing the movements of 1492 visual and telescopic meteor trails, finds proof of winds with different velocities and directions along a given trail, and both vertical and horizontal components within quite small volumes.

* Communicated by Dr. A. C. B. Lovell.

† Now at Canterbury University College, New Zealand.

Based on these observations, a trail of meteor ionization, initially linear, may be subject to one or more of the following types of motion :—

(i) Drift or tilt of the trail as a complete unit ;

(ii) Gradual curvature of large sections of the trail ;

(iii) Relative movement of short sections of the trail, due to turbulence.

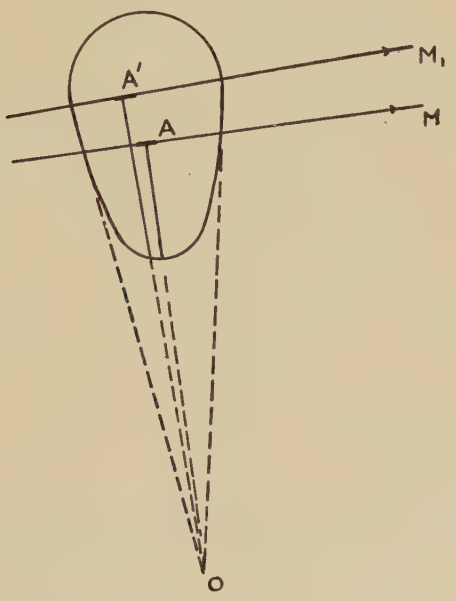
These effects will be considered in turn.

§ 3. Interpretation of Observations in Terms of Wind Motion.

(i) *Range Drift—Drift or tilt of the trail as a complete unit.*

It is shown by the work of Lovell, Banwell and Clegg (1947), and by Hey and Stewart (1947) that the reflection from a meteor trail is markedly aspect sensitive at 72 Mc./s. In fig. 1, if M is the line of motion of the

Fig. 1.



meteor, most of the echo energy received at O will come from the region close to the foot A of the perpendicular from O to the trail. The elliptical curve represents the coverage of the aerial over a horizontal plane at a height of 100 km. The meteor trails which give echoes are contained completely within this area*.

If the trail drifts as a unit nearly parallel to itself to a position M_1, under the influence of winds, the reflecting region will move to A′. The velocity of such movement should be of the order of 50 m./sec., and could

* The polar diagram has a volume distribution for meteor collection (Clegg 1948), but for present illustrative purposes the section of this volume is sufficient.

be either towards or away from the receiver. The change of range in drifts of this order is very slight, being only 1·25 km. in 25 sec. Drifts of this magnitude have been found by McKinley and Millman (1949) who report that the majority of long-enduring discrete range echoes exhibit a drift in slant range at the rate of approximately 1·0 km. in 25 sec. Their result is in good numerical agreement with the hypothesis of a general wind drift. The further observation by McKinley and Millman that the drift is predominantly away from the observer still awaits explanation.

(ii) *Range Drift—Gradual curvature of large sections of the trail.*

In addition to this slow drift, much greater range movements have been seen in a proportion of echoes. The available experimental evidence on these gross range drifts observed on 72–73 Mc./s. is given in Table I.

TABLE I.

Observations on Range Drift.

Equipment				Range drifts			
Author	Freq. (Mc./s.)	Aerial array	Beam width	Total No. of echoes examined	No. of travelling echoes	Order and sign of movement	Velocity of movement
Prentice, Lovell, and Banwell 1947	72	*Vertical* Yagi	Main beam 21°. Appreciable side lobes	1836	55·0 (3%)	5 to 10 km. Mostly decreasing range	Widely spread from a few km./sec. to 60 km./sec.
Hey, and Stewart 1947	73	*Vertical* Yagi	15°	—	2·0%	Mostly decreasing range	0 to 25 km./sec.
Ellyett *	72	Array of 5 Yagis at 15° elevation	Az. 6·5° El. 14°	>1000	(0·1%)	+7·0 km.	+0·39 km./sec.

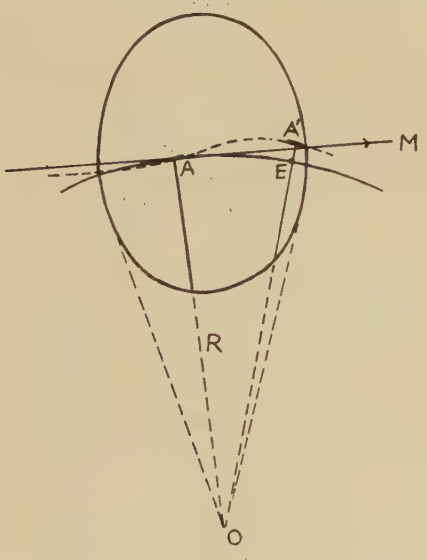
* Unpublished. From an analysis of more than 1000 cine records obtained during the major showers of 1946–48 using the technique described by Lovell, Banwell and Clegg (1947).

The first point of interest emerging from the table is the agreement between the three independent sets of observations with respect to the magnitude and the velocity of apparent movements, which are of a very high order ; speeds of 25 km./sec. being by no means exceptional. The suggestion was made by Hey and Stewart (1947) that since these drift velocities were far in excess of any possible direct air movement, the distortion of the trail must have produced a change in the point of reflection, leading to the very high apparent rate of movement. Type (ii) of the classes of air motion suggested is identical with this mechanism. Referring to fig. 2 the initial reflecting region of the trail at A is shifted to A′ by progressive distortion. The actual displacement of the trail is exaggerated in the figure, the region A′ being visualized as lying very close

to its original position along M. The one definite range drift obtained from an analysis of the cine records (Table I.) is the only echo available with sufficient information to allow an estimate to be made of the wind gradient. As the range R=167 km. and the drift EA′=+7 km. in 18·0 sec., the angular displacement EOA is 16°. This yields a wind velocity gradient of 9 m./sec./km. This value, being less than one-fifth of the mean wind velocity, appears not unreasonable.

On the mechanism of fig. 2, since reflection effectively comes from some new point along the line M, the range always increases, and must do so in all cases where the normal to a linear trail lies within the beam area for initial reflection.

Fig. 2.



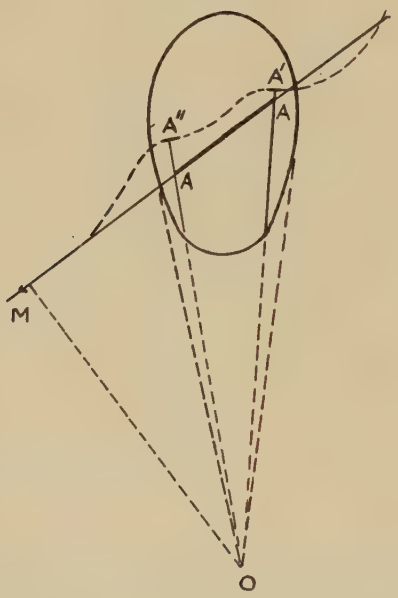
Although a percentage of outward drifts do occur, the major group appears to be directed towards the observer. To account for this group a slightly different mechanism must be postulated. Consider a meteor track, given by the heavy section AA in fig. 3, passing obliquely through the aerial beam, so that the projected perpendicular reflecting point lies outside the beam angle. In such a case no echo is received on 72 Mc./s. unless the track undergoes distortion and presents a suitable directed portion A′ within the beam. As the bending increases A′ may decrease in range towards A″.

An interesting feature of the results given in Table I. is the marked difference between the numbers of echoes showing range drift with vertically and horizontally directed beams. This can also be explained on the concept of trail distortion. With a vertically directed beam an echo can only be obtained by specular reflection if the meteor trail is

almost horizontal, or if the meteor emanates from a radiant close to the horizon. Under these circumstances the rate will be low (Prentice 1948). Hence a larger proportion of the echoes observed with a vertical beam are due to trail distortion. A simple arithmetical estimate, taking into account the relative collecting areas (Clegg 1948) shows that on this hypothesis, the percentage of range drifts observed with a narrow vertical beam should exceed that observed with a horizontal beam by a factor of the order of 20. This agrees well with the ratio obtained experimentally.

As the radio wavelength is lowered the echo duration increases, and hence the occurrence of reflection at angles other than the normal to the initial trail should increase. This has been substantiated by the work of McKinley and Millman at 32·7 Mc./s., who find that trails occurring

Fig. 3.



visually within $\pm 10^\circ$ of a direction perpendicular to the line of sight from the station to the trail produce radio echoes after a delay of about 1·0 sec., whereas for angles of departure lying between 50° and 90° the delay increases to 10·0 sec.

(iii) *Turbulent Motion—Relative movement of short sections of the trail.*

The third process considered in § 2 concerned the relative movements of short sections of the trail. The explanation of the high speed amplitude fluctuations in terms of this type of trail movement has been discussed by Greenhow (1950). In the presence of such turbulence the aspect sensitivity of the trail for radio reflections should slowly vanish, and appreciable

range spreading of long duration echoes should gradually occur, being more marked on the lower radio frequencies. Evidence for this effect has been obtained on 72 Mc./s. (Prentice, Lovell and Banwell 1947), while McKinley and Millman 1949, have found marked evidence of the effect on 32·7 Mc./s.

§ 4. Correlation of Meteor Trail Fluctuations with Direct Observations of Ionospheric Fading.

Studies of the fading of radio waves returned from the ionosphere suggest the presence of moving ionospheric irregularities (Ratcliffe 1948). These irregularities are considered to be caused by wind motion (Pawsey 1935, Findlay 1949, Mitra 1949). These measurements have been made at frequencies ranging from 16 Kc./s. to 6 Mc./s., and at such frequencies meteor ionization from individual trails will remain an effective echo source for periods of many minutes.

Table II.

Period of fading due to Relative Movement of Ionized Air at Heights above 80 km.

Method	Author	Period of fading	Equivalent period at 72 Mc./s. Short	Long
Ionospheric echoes				
Amplitude fading curves	Ratcliffe (1949)	5/min. at 3 Mc./s. (approximate mean from spread of results)		0·7 sec.
Phase fading	Findlay (1949)	1/sec. at 2·4 Mc./s.	0·03 sec.	
Phase and amplitude fading	Bracewell (1948)	0·2/min. at 16 Kc./s. 0·67/hr. at 16 Kc./s.	0·07 sec.	1·2 sec.
Meteor trail echoes				
Amplitude flutter	Greenhow (1950)	0·04 sec. at 72 Mc./s. 0·09 sec. at 36 Mc./s.	0·04 sec. 0·04 sec.	0·5 sec.
Longer period fading	Greenhow (1950)	1·0 sec. at 36 Mc./s.		
Amplitude flutter	McKinley and Millman (1949)	5–10/sec. at 32·7 Mc./s.	0·09–0·05 sec.	

It has already been suggested by several authors that sporadic E effects are in part due to meteor ionization (summary by Lovell 1948). Evidence will now be produced which indicates that some of the irregularities causing the fading of radio waves from the ionosphere may also be due directly to wind distortion of meteor trails.

Since no measured wind velocities are yet available from meteor studies, the mean period of fading must be used as the point of comparison. In order to allow comparison of the results of various workers, it will be assumed that the speed of fading is proportional to the radio frequency employed. This assumption has already been made by Ratcliffe (1948),

who has found some evidence of its validity. Further experimental evidence supporting this assumption has now been obtained by observation of the fluctuations occurring in the amplitude of radio waves reflected from meteor trails on 36 and 72 Mc./s. sumultaneously (Greenhow 1950).

Two distinct periods of ionospheric fading have been recognized on the very long wavelengths (Bracewell 1948) and although a large scaling up in frequency is required it is remarkable that the figures at 72 Mc./s. agree closely both with the short period flutter and with the longer period observed from meteor trails. A recent figure quoted by Ratcliffe at intermediate frequencies agrees with the longer period, and the figure given by Findlay (1949) is almost identical with the meteor flutter period.

There appears, then, to be a close correspondence in the flutter observed from meteor trails and the ionospheric fading obtained from unspecified " centres ". Both show similar characteristics and have the same mean period. It therefore seems not unlikely that some of the ionospheric fading is, in fact, directly due to meteor ionization at E region heights.

In the absence of further results, no great significance can as yet be attached to the correspondence of the longer periods.

Acknowledgments.

The work described in this paper was carried out at the Jodrell Bank Experimental Station of the University of Manchester. Valuable advice was given by Dr. J. A. Clegg, and particularly by Dr. A. C. B. Lovell, Director of the station. The author is personally indebted to the Imperial Chemical Industries, Ltd., for the award of a Research Fellowship, and to the Council of Canterbury University College, Christchurch, New Zealand, for the granting of three years leave of absence.

References.

Bracewell, R. N., 1948, *Nature, Lond.*, **162,** 9.
Clegg, J. A., 1948, *Phil. Mag.*, **39,** 577.
Findlay, J. W., 1949, Private Communication.
Greenhow, J. S., 1950, *Phil. Mag.*, **41,** 682.
Herlofson, N., 1947, *Nature, Lond.*, **160,** 74.
Hey, J. S., and Stewart, G. S., 1947, *Proc. Phys. Soc.*, **59,** 858.
Lovell, A. C. B., Banwell, C. J., and Clegg, J. A., 1947, *Mon. Not. R. Astr. Soc.*, **107,** 164.
Lovell, A. C. B., 1948, *Rep. Prog. Phys.*, **11,** 415 (London : Physical Society).
McKinley, D. W. R., and Millman, P. M., 1949, *Proc. Inst. Radio Engrs.*, **37,** 364.
Mitra, S. N., 1949, *Proc. Inst. Elect. Engrs.*, Pt. III, **96,** 441.
Olivier, C. P., 1947, *Proc. Amer. Phil. Soc.*, **91,** 315.
Pawsey, L. J., 1935, *Proc. Camb. Phil. Soc.*, **31,** 125.
Prentice, J. P. M., Lovell, A. C. B., and Banwell, C. J., 1947, *Mon. Not. R. Astr. Soc.*, **107,** 155.
Prentice, J. P. M., 1948, *Rep. Prog. Phys.*, **11,** 389 (London : Physical Society).
Ratcliffe, J. A., 1948, *Nature, Lond.*, **162,** 9 ; 1949, Private communication.
Störmer, C., 1939, *Astrophysica Norwegica*, **3,** 117.

LXIII. *Nuclear Transmutations Produced by Cosmic-Ray Particles of Great Energy.*—Part V. *The Neutral Mesons.*

By A. G. Carlson *, J. E. Hooper and D. T. King,
The H. H. Wills Physical Laboratory, University of Bristol †.

[Received May 24, 1950.]

[Plate XXIV.]

Summary.

The spectrum of the γ-radiation in the atmosphere at 70,000 ft. has been determined by observations on the scattering of pairs of fast electrons recorded in photographic emulsions exposed in high-flying balloons. The detailed form of the spectrum is consistent with the assumption that the γ-rays originate by the decay of neutral mesons. It is found that the mass of the neutral mesons is 295 ± 20 m_e, and that they are created in nuclear explosions with an "energy spectrum" similar to that of the charged π-particles.

A method is described for determining the lifetime of the neutral mesons and their frequency of occurrence compared with charged π-particles. The lifetime, τ_{π^0}, is less than 5×10^{-14} sec. It may be possible to determine the lifetime, by observations of greater statistical weight, if it is longer than 2×10^{-14} sec. The ratio of the number of neutral mesons to charged mesons, produced in nuclear explosions of great energy, is equal to $0{\cdot}45 \pm 0{\cdot}10$.

1. Introduction.

Recent experiments by Bjorklund, Crandall, Moyer and York (1950), strongly suggest that the high energy γ-rays observed when matter is bombarded by protons with an energy greater than about 200 MeV. are of secondary origin; that they arise by the decay into two quanta of neutral mesons, of mass about 300 m_e and lifetime less than 10^{-11} secs., which are the primary products of the nuclear collisions. Further, Bradt, Kaplon and Peters (1950), have made a detailed analysis of a disintegration produced in a photographic emulsion by an α-particle with an energy between 10^{12} and 10^{13} eV., which leads to the emission of fifty-seven shower particles. In the "core" of the shower particles, most of which are mesons (Piccioni 1950, Fowler 1950), these authors observe the production of several pairs of electrons, and conclude that the charged mesons are accompanied by γ-radiation. If this radiation is also created by the decay of neutral mesons—and if the latter particles are produced in the nuclear explosion with values of the kinetic energy similar to those of the charged mesons—it may be shown that the proper lifetime of the neutral mesons is less than 3×10^{-13} secs.

* On leave from the University of Upsala.
† Communicated by Professor C. F. Powell, F.R.S.

Finally, the existence of neutral mesons of mass about 280 m_e, and their mode of decay into two γ-rays, appears to have been put beyond doubt as a result of experiments by Panofsky, Aamodt and York (1950) on the capture of π^--particles in liquid hydrogen. This process leads to transmutations which may be represented by the following equations :—

$$\text{(i)}\quad H^1+\pi^{-1} \rightarrow n^0+\pi^0\,; \qquad \pi^0 \rightarrow 2h\nu \ (h\nu \sim 70 \text{ MeV.})\,;$$

and

$$\text{(ii)}\quad H^1+\pi^{-1} \rightarrow n^0+h\nu \ (h\nu \sim 140 \text{ MeV.})\,;$$

π^0 representing a neutral meson.

The observed degree of homogeneity of the γ-rays from reaction (i), which is the more probable, indicates that the π^0-particles are emitted with small kinetic energy. This observation, and the known masses of the other particles involved in the transmutations, proves that the mass of the particle π^0 is only a few electron masses less than that of the charged π-particles ; viz. $m_{\pi^0} \sim 280\ m_e$.

These observations appear to have an important bearing on experiments carried out in the last few years with Wilson chambers by Cocconi *et al.* (1946), Fretter (1949), Chao (1949), Cocconi (1949), and Gregory *et al.* (1950). These authors have shown conclusively that the cascade showers of the soft component are commonly associated with penetrating showers of charged mesons.

2. Scope of the Present Experiments.

In Section I. of the present paper, observations are described which show that the detailed features of the spectrum of the γ-radiation accompanying the penetrating showers are consistent with the assumption that the quanta are created by the decay of neutral mesons of mass $295 \pm 20\ m_e$; and that the distribution in energy of the neutral particles is closely similar to that of the charged mesons of the " showers ".

In Section II., a method of determining the lifetime, τ_{π^0}, of the neutral mesons is described, and the results prove that $\tau_{\pi^0} < 5 \times 10^{-14}$ sec. There is some indication that $\tau_{\pi^0} = 3 \times 10^{-14}$ sec. ; and if this conclusion can be confirmed by further measurements, the observations would establish the independent existence of neutral mesons. Finally, the results show that the relative frequency of production of the neutral and charged mesons, $\mathcal{N}(\pi^0)/\mathcal{N}(\pi^\pm) = 0{\cdot}45 \pm 0{\cdot}1$. It appears, therefore, that on the average, one neutral meson is produced for every two charged π-particles in the explosive disintegrations which give rise to the penetrating showers. This result, taken in conjunction with observations by Rossi on the relative amounts of energy carried by the " hard " and " soft " components of the cosmic radiation in the atmosphere, shows that the latter arises largely as a result of nuclear interactions, but indirectly through the production of neutral mesons ; and that at 70,000 ft., the fraction which need be attributed to other processes such as the decay of μ-mesons, " knock-on " electrons, etc. is negligible..

Section I.

Determinations of the Mass and the Distribution in Energy of Neutral Mesons.

3. The γ-Ray Spectrum at 70,000 ft.

During the past year observations have been made in this laboratory on photographic plates exposed at an altitude of 70,000 ft.; see Camerini *et al.* (1949). A number of these plates have been examined under the microscope with relatively high magnification, in order to detect pairs of electrons and other events involving only particles of charge e moving at relativistic velocities. Such events commonly escape observation when the plates are searched at lower magnification. By determining the energy of the pairs of electrons by the scattering method, in those favourable cases in which the tracks in the emulsion are of sufficient length, it has been possible to deduce the energy spectrum and direction of motion of the photons of the "soft" component at 70,000 ft.

Table I.

Energy region (MeV.)	Cell length (microns)	Probable error in energy (per cent)
10–100	100	19
100–200	200	28
200–400	250	32

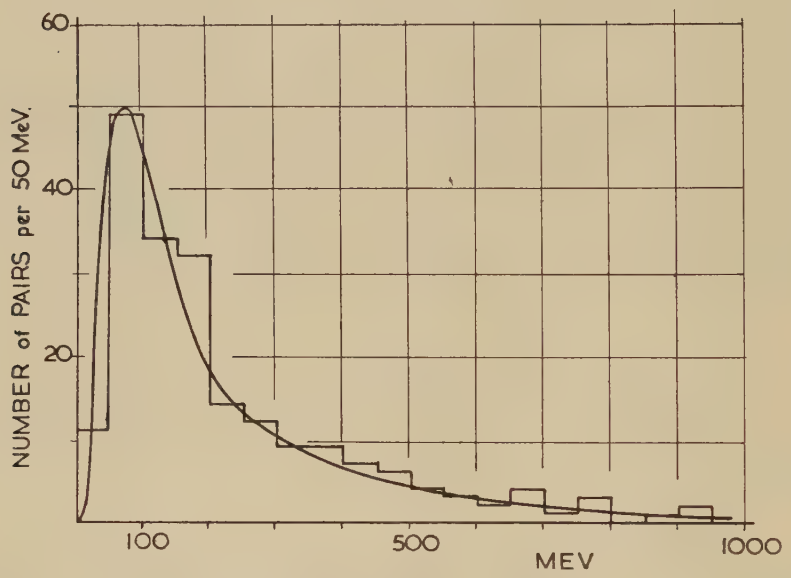
The above errors are given on the assumption that the track under consideration is just 1000 microns long.

The method employed in the determination of the energy of the electrons was similar to that described by Fowler (1950) and King (1950). Only those pairs were accepted for measurement of which the track of each component was longer than 1 mm. In the case of tracks of length less than 2000 μ, it is not possible to determine the energy of the corresponding particles unless it is less than 500 MeV. On the other hand, satisfactory estimates can be made up to 400 MeV. with tracks only 1000 μ long. Characteristic values of the probable errors of the measurements, for tracks 1000 μ long, are shown in Table I. The energy of a pair is usually substantially greater than that of one of its components. We may, therefore, say that the total energy, E, of a pair may be measured with a probable error of less than 30 per cent, for values of $E < 600$ MeV., if both tracks are longer than 1000 μ; and for $E < 1500$ MeV., if the tracks are longer than 2000 μ.

The results of measurements on pairs of very low energy require a small correction to allow for the influence of scattering on the escape of particles from the emulsion. Such corrections are, however, very small for $E > 20$ MeV., and they have not been applied in making the present measurements. The results are shown in fig. 1.

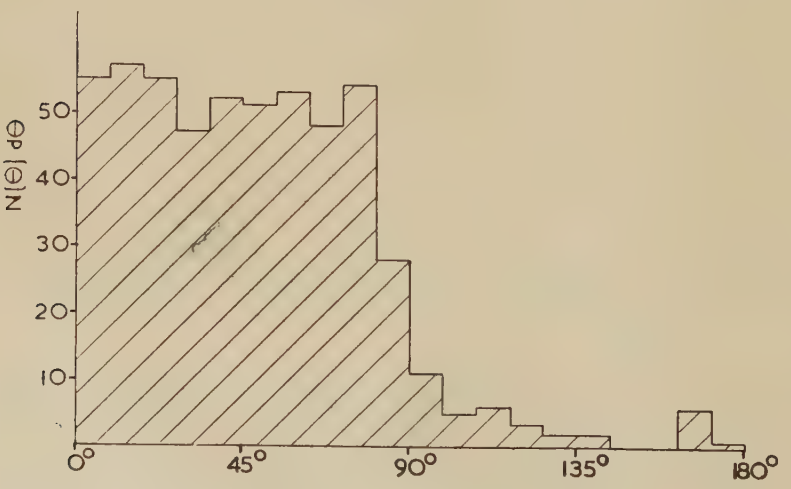
In determining the orientation of the pairs of electrons with respect to the vertical line through the plates during the exposure, it was considered sufficient to determine only the inclination, θ, of the projection of the "bisector" of the pair. For pairs of electrons produced by radiation with

Fig. 1.



Distribution in energy of the pairs of electrons observed at 70,000 ft. The full line shows the expected photon spectrum arising by the decay of neutral mesons of mass 280 m_e and with kinetic energies similar to those of charged π-particles emitted from stars.

Fig. 2.



Angular distribution of the directions of motion of photons at 70,000 ft.

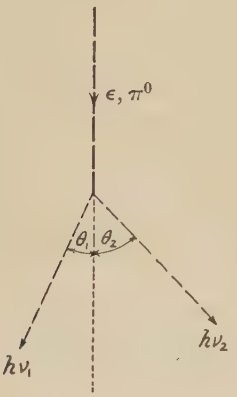
values of the quantum energy in the range under consideration, the " bisector " gives a measure of the direction of the parent γ-ray which is subject to a probable error of less than 0·2°.

The distribution in the values of θ, deduced from observations of 500 pairs is shown in fig. 2. It will be seen that the directions of motion of the quanta are distributed nearly isotropically in the range of angles from 0° to 90° with the direction vertically downwards, and that the intensity of any upward directed radiation is relatively very small. Charged mesons ejected in nuclear disintegrations in this exposure exhibit the same angular distribution; Davies, Franzinetti, and Perkins (1950).

4. Characteristics of the Radiation Produced by the Decay of Neutral Mesons.

Following the publication of the results of Bjorklund *et al.*, it occurred to us that if the γ-radiation in the atmosphere arises by the decay of neutral mesons—and if it is examined in conditions in which it has not been seriously modified by " bremsstrahlung " and other processes associated

Fig. 3.



with the development of the cascade showers of the soft component— its spectrum should display certain features of an internal consistency, and should allow a determination to be made of the mass of the postulated neutral meson. The method depends on the following considerations :—

Suppose a neutral particle of mass m_0, moving with velocity βc, transforms into two quanta, of energy $h\nu_1$ and $h\nu_2$ emitted in directions making angles θ_1 and θ_2 with the line of motion of the parent particle—see fig. 3. The equations for the conservation of energy and of momemtum may be written :

$$\frac{h\nu_1}{c}\sin\theta_1=\frac{h\nu_2}{c}\sin\theta_2\,; \qquad \frac{h\nu_1}{c}\cos\theta_1+\frac{h\nu_2}{c}\cos\theta_2=m_0\beta c/\sqrt{(1-\beta^2)}\,;$$

$$h\nu_1+h\nu_2=m_0c^2/\sqrt{(1-\beta^2)}=\mathrm{B}m_0c^2=\epsilon.$$

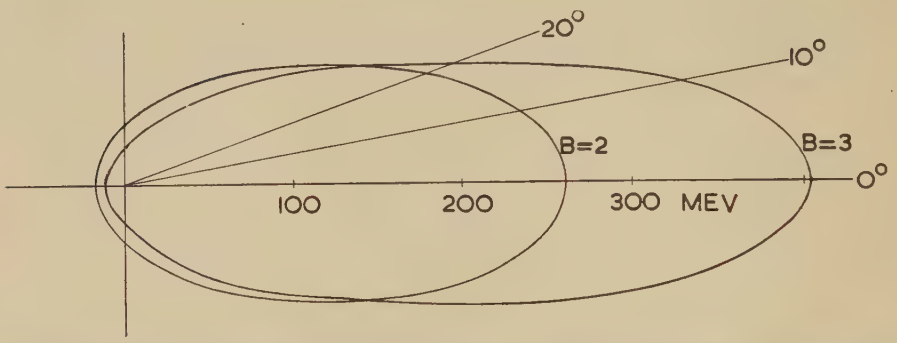
It is then easily shown that the energy of the quantum emitted at an angle θ is given by the well known formula for relativistic Doppler effect,

$$h\nu = \frac{m_0c^2}{2B(1-\beta\cos\theta)}. \quad \ldots\ldots\ldots \quad (1)$$

Characteristic curves showing the variation of $h\nu$ with θ, for different values of B and taking m_0c^2 equal to 140 MeV., are given in fig. 4. It follows from (1) that for a given value of β, the maximum and minimum values of $h\nu$ are given by the relations

$$h\nu\,(\max) = \frac{Bm_0c^2}{2}(1+\beta)\,; \qquad h\nu\,(\min) = \frac{Bm_0c^2}{2}(1-\beta)\,;$$

Fig. 4.



Polar diagram showing the variation of energy with the direction of emission of the photons resulting from the decay of neutral mesons. Calculations for two values of the total energy of the neutral mesons.

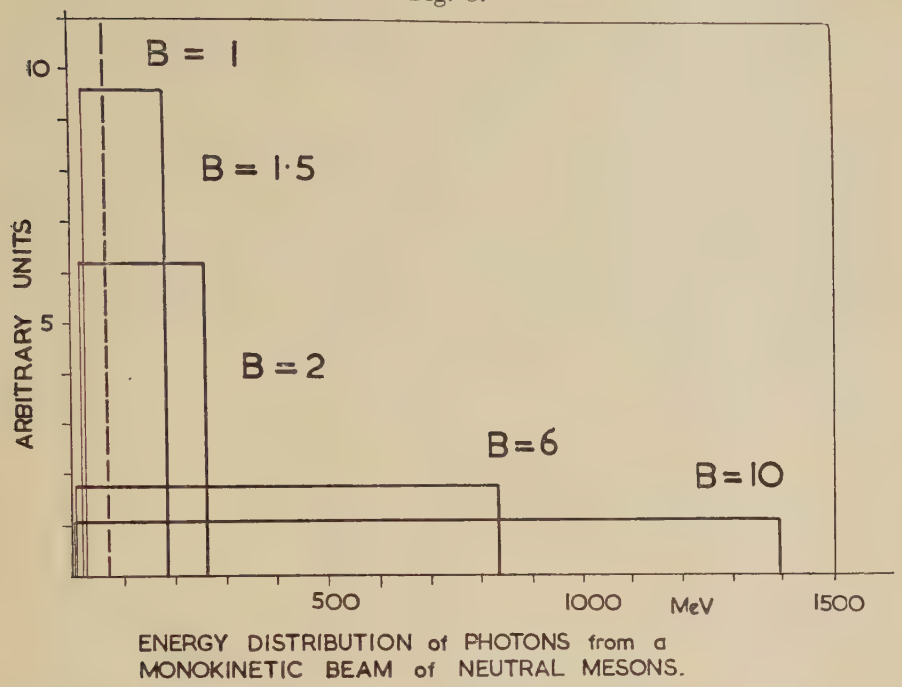
and that the greater the velocity of a homogeneous beam of neutral mesons, the greater the width of the spectrum of the emitted γ-radiation. Assuming that the γ-radiation is emitted isotropically in a coordinate system moving at a velocity equal to that of the mesons, the form of the spectrum of the γ-rays can be calculated—see Appendix—and can be shown to be given by the radiation relation

$$N(E)\,.\,dE = k\,.\,dE,$$

where $N(E)\,.\,dE$ is the number of emitted quanta of energy between E and $E+dE$. k is a constant for values of E lying between $h\nu$ (max) and $h\nu$ (min); and $k=0$ for values of E outside this interval. Characteristic spectra for different values of B, normalized to correspond to the same total number of quanta, are shown in fig. 5.

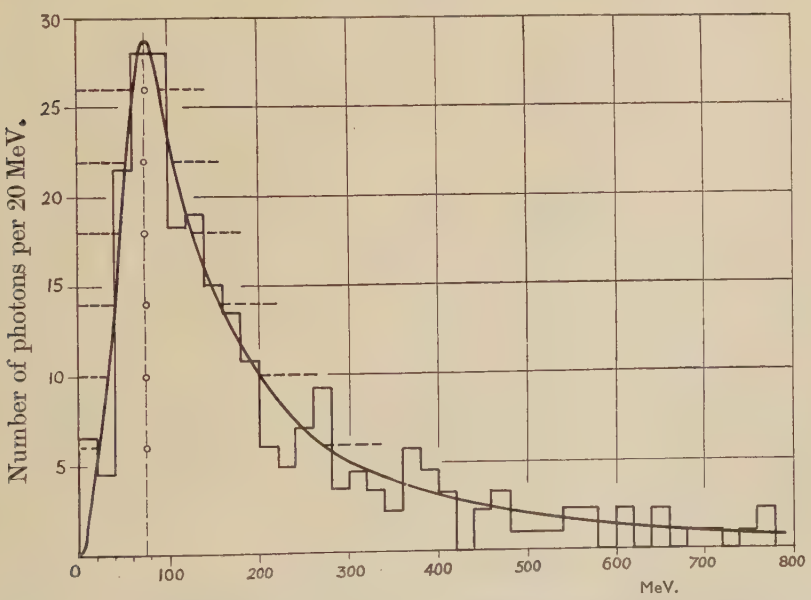
In the case of an inhomogeneous beam of neutral mesons, dispersed in velocity, the resulting γ-ray spectrum can be shown to have the features illustrated in fig. 6. There is a cusp in the spectral distribution if a finite number of neutral mesons are emitted with zero velocity. For any value

Fig. 5.



Distribution in energy of the photons produced by the decay of neutral mesons.

Fig. 6.



Spectrum of the γ-radiation at 70,000 ft. The full line is that giving the best fit with the experimental results, and the mass of the neutral particle has been deduced from it.

of the intensity below the maximum, there are two corresponding values of the quantum energy E_1 and E_2, and these can be shown to be related to the rest-mass m_0 of the parent neutral particles by the relation

$$\sqrt{(E_1E_2)}=m_0c^2/2. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

The proof of this theorem is included in the Appendix.

It follows from the above result that if most of the γ-radiation at 70,000 ft. is produced by the decay of neutral mesons—and if it is not seriously modified by the processes which lead to the formation of the cascade showers—there should be an inner consistency in the form of the spectrum corresponding to the relation displayed in equation (2). It appeared to us that the observation of such a consistency would give very powerful support for the view that the γ-radiation does in fact arise by a decay of neutral mesons and would allow an estimate to be made of their mass.

TABLE II.

Measurements of the Neutral Meson Mass from the Observed Energy Distribution of the Photons.

Weights (No. of pairs)	E_1	E_2	$\epsilon_0/2=\sqrt{(E_1E_2)}$	Weighted $\epsilon_0/2$
26	63	91	75·5	196
22	53	106	75·0	165
18	45	124	75·0	135
14	39	156	78·0	109
10	31	200	79·0	79
6	21	274	75·5	42

Total 96 Weighted mean for $\epsilon_0/2=75{\cdot}8$ MeV.

All energies are given in MeV. Mass of neutral particle $m_{\pi^0}=295\pm20m_e$.

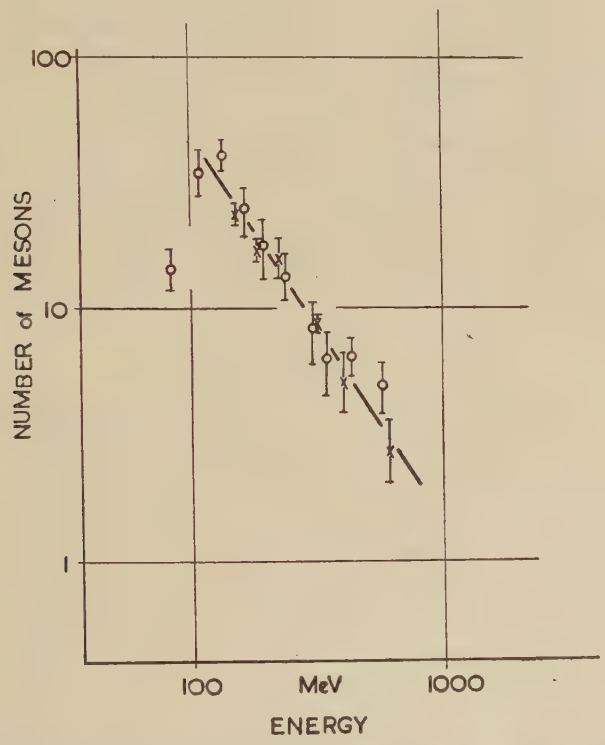
5. COMPARISON WITH EXPERIMENT ; THE MASS OF THE NEUTRAL MESON.

Fig. 6 shows the distribution in energy of the γ-radiation. It has been deduced from that of the pairs (see fig. 1) by assuming a variation with energy of the cross-section for pair production given by the quantum electro-dynamics (Heitler 1944). The full line in best accord with the distribution has been drawn, and intercepts made at various arbitrarily chosen intensities to determine values of E_1 and E_2. The corresponding values of $\sqrt{(E_1E_2)}$ are shown by circles. It will be seen from fig. 6 and Table II. that they show a remarkable degree of consistency, and indicate that the γ-radiation arises by the decay of neutral mesons of mass $295\pm20\ m_e$. The results may be subject to a small systematic error arising from a variation with energy of the scattering constant k, defined by the equation $E=k/\bar{\alpha}$, where $\bar{\alpha}$ is the mean observed angle of scattering for a cell length of 100 μ. Such a variation may exist as a result of the effects of screening.

The stated errors in the estimated mass of the neutral meson correspond only to those due to statistical fluctuations in the numbers of measured pairs of electrons, and when the result was first obtained it was not clear to what extent it was influenced by the development of the cascade showers. The recent observations of Panofsky *et al.* (1950) suggest, however, that the " distortion " in the form of the spectrum is small, and this appears to be associated with the following features of the experiments.

Firstly, the plates were exposed near the main meson production layer in the atmosphere. At this altitude the probability that a given quantum will have produced a cascade, even in the case of quanta produced at the

Fig. 7.



The distribution in energy of the neutral mesons emitted from nuclear disintegrations. Spectrum of neutral mesons shown thus : × ; and of charged mesons, as observed by Camerini *et al.* (1950), shown thus : ○.

top of the atmosphere, is small. Secondly, if the cascade process starts in the plate assembly itself, it will commonly give rise to relatively narrow showers, made up of a number of particles moving nearly parallel to one another and close together. The development of such cascade showers has been observed in the present experiments, but they are rare, and any pairs associated with them have been excluded from the analysis.

For these reasons we regard our observations as giving very strong evidence for the view that a large part of the γ-radiation in the atmosphere is due to the production and decay of neutral mesons, of mass $\sim 295 \pm 20 m_e$, which may be identified with the particles called for by the results of experiments by Bjorklund *et al.* (1949), and of Panofsky *et al.* (1950).

6. Distribution in Energy of the Neutral Mesons.

From the observed form of the γ-ray spectrum it is possible to deduce the distribution in energy of the parent neutral mesons and the result, determined by the methods described in the Appendix is shown in fig. 7. The corresponding distribution for the charged mesons, emerging from the nuclear explosions recorded in the same plates and determined by Camerini *et al.* (1950 a), is included in fig. 7, and it will be seen that the two "spectra" are similar in form. Alternatively, the spectrum of the γ-rays which would be produced by the decay of neutral mesons, of mass 280 m_e and with an energy distribution similar to that of the charged mesons, can be easily computed. Using the results of Camerini *et al.*, for the charged mesons, the result thus obtained is shown by the smooth line in fig. 1. It will be seen that the calculated form of the spectrum is indistinguishable from that observed, and added weight is thus given to the basic assumption regarding the origin of the γ-rays, and for the view that the parent neutral mesons are created with a distribution in energy closely similar to that of the charged particles.

Section II.

Observations to Establish the Independent Existence of Neutral Mesons, and to Measure their Lifetime and Frequency of Occurrence.

7. A Method of Determining the Lifetime.

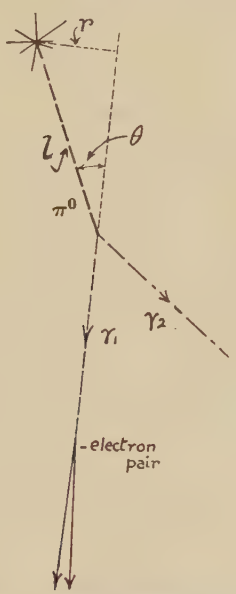
The evidence for the existence of neutral mesons presented hitherto, whilst very strong, is circumstantial and depends on observing γ-radiation of which the detailed characteristics are consistent with an assumed origin by the decay of neutral mesons. Although the mass of evidence which has now been accumulated is so large that the conclusion to which it leads can hardly be doubted, it is nevertheless important to obtain an unambiguous proof of the independent existence of the particles. Further, it has been pointed out by Fukuda and Miyamoto (1949) that the determination of the lifetime of the neutral meson is of great importance for distinguishing between different theoretical treatments of the nature of the particle. The following considerations indicate that such evidence could be obtained by an application of the photographic method if the lifetime of the particles is greater than about 2×10^{-14} sec.

Suppose that a neutral meson emerges, with a total energy $\epsilon = B m_0 c^2$, from a nuclear explosion occurring in the emulsion ; and that it decays into two photons, after traversing a distance l. We have seen that the two γ-rays will in general be emitted in directions inclined to that of the

parent particle. Suppose, further, that one of the photons produces a pair of electrons in the emulsion within a distance of 1 or 2 mm. of its point of creation. The probability of such an occurrence is about 3 per cent for each emitted neutral meson ; for the average length of path in the emulsion of a γ-ray before pair production is 46 mm. *, there are two photons produced by each neutral particle, and a considerable proportion of the events will escape observation because of the finite thickness of the emulsion.

We have seen that for neutral mesons with kinetic energies in the range of values with which we are concerned, the pair of electrons will move in directions inclined to one another at an angle of less than 0·2°.

Fig. 8.



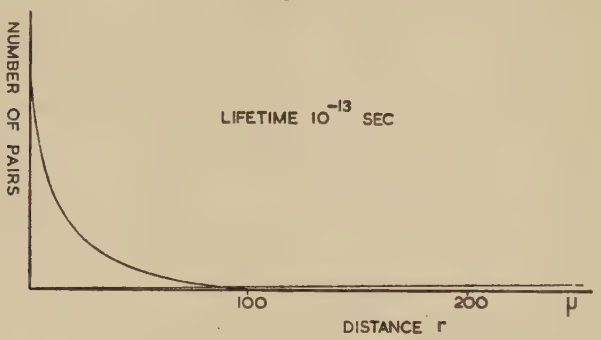
Principle of the method for the determination of the lifetime of the neutral meson.

Apart from the momentum imparted to the nucleus involved in the formation of the pair (Bethe 1934), the bisector of the angle between the tracks of the two electrons therefore gives a measure of the direction of motion of the parent γ-ray which is subject to errors of the order of 0·2° (see fig. 8). It follows that if the mean line of the pair is produced backwards, it will not pass precisely through the disintegrating nucleus in which the neutral meson originated, but at a distance r from it. The greater the value of the lifetime, the greater the average distance l and the corresponding values of r.

* This value has been calculated by taking account of the observed energy spectrum of the γ-radiation, and the variation of the conversion length with quantum energy.

The value of r for any particular event will depend upon the distance travelled by the neutral meson before it decays, and the direction of emission of the photon relative to the line of motion of the parent particle. For a homogeneous beam of neutral mesons, moving with a given energy, the expected distribution in the values of r can be calculated by the methods outlined in the Appendix. From the results thus obtained, it is then possible to proceed to the general case in which the neutral mesons have a distribution in energy. Fig. 9 represents the results of such computations, assuming an energy distribution similar to that of the charged mesons deduced by Camerini *et al.* (1950 a), and a lifetime of the neutral mesons equal to 10^{-13} sec.

Fig. 9.



The calculated distribution in the values of the quantity r for an assumed lifetime of the neutral meson of 10^{-13} sec., together with the calculated form of the distribution due to unrelated pairs.

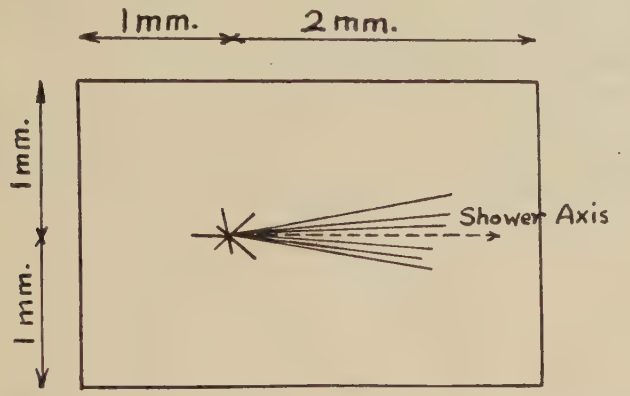
It is a particular feature of the method that in the case of mesons of high energy, $B \geqslant 2$, the calculated distributions are almost independent of the value of B, and of the distribution in energy of the neutral particles. This is due to the fact that although the emitted radiation tends to be closely collimated about the line of motion of the parent mesons as B increases—a factor which leads to a reduction in the values of r—the effect is compensated by the relativistic extension of the time–scale of the moving particles. This tends to make the neutral mesons survive over a longer length of path in the emulsion.

8. Experimental Results.

In order to apply the above method an extensive search has been made at high magnification of the region of the emulsion in the immediate neighbourhood of nuclear explosions occurring in the plates in which three or more " shower " particles are created ($n_s \geqslant 3$). All pairs of electrons observed in the volume of the emulsion defined in the manner illustrated in fig. 10 were recorded, and the corresponding values of r determined by observations on the direction of the projection of the mean line of the pair

in relation to the star, and observations of " dip ". The methods are discussed in more detail in a later paragraph with special reference to the errors of measurement.

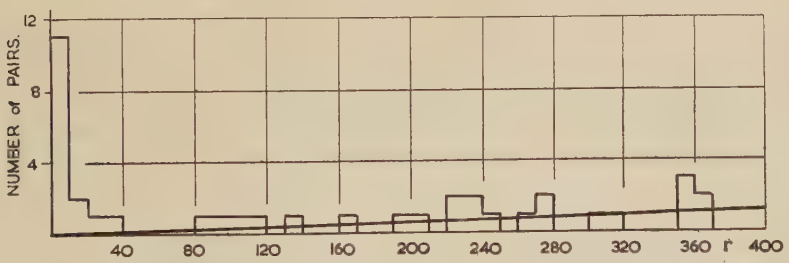
Fig. 10.



The area around a " shower ", which is scrutinized for " pairs ".

The observed distribution in the measured values of r is given in fig. 11, and displays a group of events corresponding to small values of r together with a pronounced general background. These features may be interpreted in the following manner:—We have seen that there is a random distribution of pairs of electrons in the plates due to the flux of γ-radiation, and these will sometimes, by chance, occur in the volume of emulsion round a star which is subject to scrutiny. The frequency of occurrence of these chance juxta-positions, which we may refer to as " unrelated " pairs, can be calculated from the observed numbers, per unit area of the plates, of the showers ($n_s \geqslant 3$), and of the pairs of electrons.

Fig. 11.



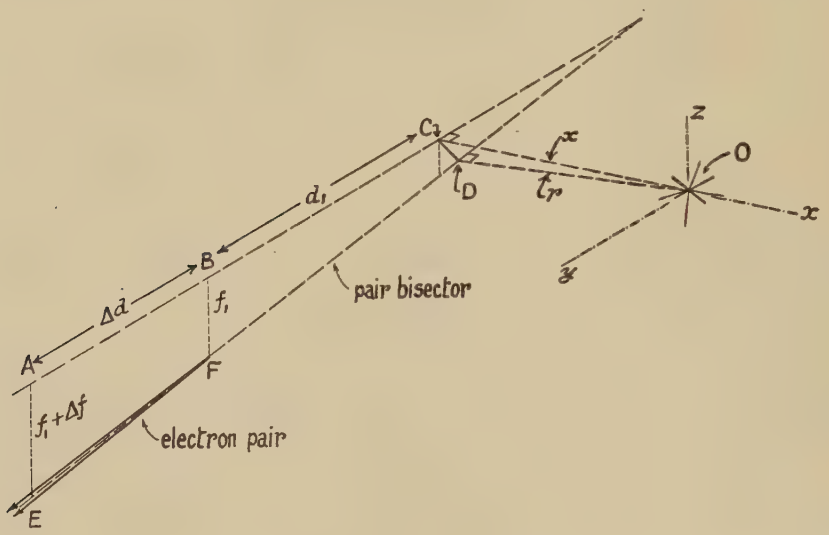
Observed distribution in the values of r for $r<400\,\mu$. The group of observations for small values of r, due to " related " pairs is clearly shown.

The calculated variation with r of the frequency of occurrence of these unrelated pairs is shown by the full line in fig. 11 ; and it will be seen to be in good accord with the observations. Thus in the interval in r from $40\,\mu$ to $200\,\mu$, the expected number of pairs is 22 ± 5 and the observed

number 23. On the other hand, most of the fifteen pairs for which $r < 40\,\mu$ must be attributed to γ-radiation associated with the neighbouring explosion. Thus the expected number of unrelated pairs for this interval in the value of r is only 0·2 compared with the 15 observed. The observations therefore prove that γ-radiation is produced, by some mechanism, as a result of nuclear interactions leading to the production of shower particles of multiplicity $n_s \geqslant 3$.

The observations on the related pairs prove that the lifetime of the neutral mesons is less than 5×10^{-14} sec., and they appear to indicate that $\tau_{\pi^0} = 3 \times 10^{-14}$ sec. It has already been emphasized that such a result, apart from its interest in establishing the decay-constant for the neutral

Fig. 12.



Determination of the quantity r.

π-particles, would also provide a direct proof for their independent existence. It is therefore important to examine what weight can be put on the observations; and, in particular, whether they can be interpreted, alternatively, as due to errors of measurement, the lifetime of the π^0-particles being much shorter than can be established by this method.

9. Errors of Measurement in the Determination of the Quantity r.

The quantity r which is taken as a measure of the relationship of a pair to a particular nuclear explosion, is illustrated in fig. 12. It is the perpendicular distance from the centre of the star to the bisector of the lines of motion of the two electrons into which the γ-ray transforms. To determine r, the perpendicular distance, x, from the star to the vertical plane containing the bisector of the electron-pair is first measured. By

" vertical " is meant here the direction normal to the plane of the emulsion which is taken as " horizontal ". x is represented by the line OC in fig. 12, and is determined by turning the plate until the bisector is parallel to one of the stage motions, Y, and coincident in the field of view with a line in the eye-piece graticule. The graticule is also provided with a scale perpendicular to the line. After moving the stage parallel to the Y axis, and refocusing, the star is eventually brought into the field of view. Its distance from the line can then be measured by means of the eye-piece scale.

The above method of determining x is subject to errors which depend on the distance BC$=d$ in fig. 12, and on the precision with which the bisector of the pair can be defined. Studies of scattering of fast particles (Davies *et al.* 1949) indicate that the alignment with a rectilinear track can be made with an error of less than 0·1°. In the case of the pairs produced by high energy photons, the two tracks are not appreciably scattered in the first 100 μ of the trajectory, and the angle between them, δ, is small ; for 100 MeV. γ-rays, $\delta \sim 0{\cdot}2°$, and for 400MeV., $\delta \sim 0{\cdot}05°$. For such energetic radiation, which constitutes most of the spectrum (see fig. 1) the definition of the bisector can be made with a precision of the order of 0·25° and the corresponding error in x can be written $\delta x \sim \frac{d}{200}$, δx and d both being measured in microns.

The second quantity necessary for the determination of r is obtained from observations on the angle of " dip " of the bisector of the pair. In fig. 12, AC is the projection of the bisector on a plane, parallel to the surface of the emulsion, which contains the centre of the star, O, and f is the depth below this plane of the point of origin of the pair, F. At a distance Δd along the bisector, at the point E, the depth is $f+\Delta f$. It is then easily shown that

$$z^2=\mathrm{CD}^2=(d\,.\,\Delta f-f\,.\,\Delta d)^2/(\Delta d^2+\Delta f^2),$$

and that

$$r^2=x^2+z^2.$$

In order to determine the probable errors resulting from the errors of measurement of the depth, numerous individual observations have been made on a number of events, and the standard deviations in the results deduced empirically. Similar observations have been made on protons of great energy which are observed to produce disintegrations, and which are known from scattering measurements to suffer an average change in direction along the trajectory of less than 0·01° per 100 μ. Values of " r " can thus be deduced for tracks which are known to pass through the centre of the star, and of which the departure from rectilinearity is negligible.

These observations indicate that the probable errors in the determination of the angle of dip are about 0·2° and that the uncertainty in the true values of r is contributed to almost equally by the two measurements,

of x and of z. According to the particular characteristics of the individual events, one or other of the two measurements may be the more important. Taking both measurements into account, the probable error in the determination of the direction of the bisector is estimated to be 0·3°.

The important question remains of the precision with which the bisector gives a measure of the line of motion of the γ-ray. In general the lines of motion of the two electrons are not equally inclined to that of the γ-ray which produces them. Further, the nucleus involved in the creation of the pair receives an impulse of order of magnitude $m_e c$. (Bethe 1934, Modesitt and Koch 1950). At the energies with which we

Table III.

Individual r measurements and their errors.

Pair	Star type	x	d	dip$=\Delta f/\Delta d$	f	r
1	7+6N	0	0	—	—	0
2	6+4p	0·5±0·5	73	0·050±0·013	2·4±0·5	1·5±1·5
3	3+5p	0·0±1·0	123	0·096±0·010	9·5±0·7	2·0±2·0
4	12+3p	2·0±2·0	497	0·120±0·010	60·0±8·0	2·0±4·0
5	9+7p	1·0±1·0	722	0·122±0·004	91·5±0·8	3·5±3·5
6	6+7p	3·5±1·0	1035	1·011±0·020	132·9±0·4	3·5±2·5
7	7+12p	1·0±1·0	378	0·094±0·004	31·8±0·5	3·5±2·0
8	21+18p	3·5±2·0	660	0·172±0·003	113·6±0·4	4·5±2·0
9	3+26p	3·0±2·0	1027	0·158±0·002	157·5±0·3	5·0±3·5
10	7+6N	6·0±1·0	226	0·337±0·013	78·6±0·5	7·0±2·0
11	3+26p	2·0±2·0	536	0·042±0·016	27·3±0·4	7·5±5·5
12	13+5p	6·0±3·0	1993	·0008±·0008	17·4±1·2	17·0±3·0
13	17+6p	7·0±1·0	395	0·320±0·006	145·0±1·1	17·5±3·0
14	14+3N	10·0±3·0	1080	0·049±0·003	81·6±0·8	29·0±4·0
15	11+4N	23·0±5·0	745	0·170±0·004	94·6±0·7	38·0±4·5

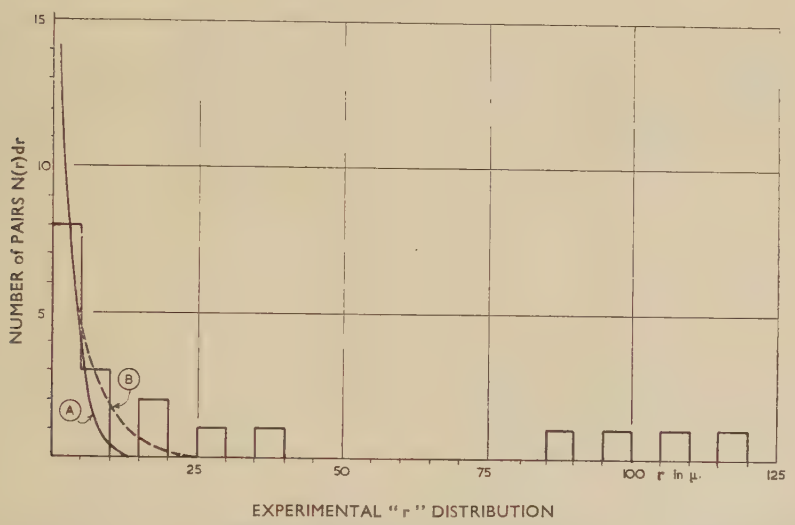
All the errors are expressed in terms of the standard deviation.
x, d, f and r are all measured in microns.

are concerned, the probable errors resulting from these effects can hardly exceed 0·2°. We thus obtain 0·4° as the final figure for the probable error in the determination of the line of motion of the photon. It follows that for errors in the measurement of the perpendicular distance from the star to this line, we may write $\delta r \lesssim d/150$.

Details of the observations on the fifteen related pairs are given in Table III. The limits of error given in the last column are those deduced from the standard deviations associated with the inconsistence in a large number of measurements on the same event. They do not, therefore, include contributions due to a deviation between the direction of the γ-ray and the bisector of the pair.

A histogram showing the values of r on a more extended scale has been reproduced in fig. 13. Included in the figure are curves indicating the distributions that would be expected for standard deviations in the errors of measurement, arising from all sources, of 0·25 and 0·5°, the lifetime of the particles being assumed to be short compared with 10^{-14} sec. It will be seen that there are only two or three pairs for which the values of r are substantially greater than those expected from a standard deviation of 0·5°. Reference to Table III. shows, however, that the values of r are many times greater, for these four events, than the standard deviations. It is very improbable that four such events should be found among a sample of fifteen related pairs, if the γ-rays diverge, in effect, from the nuclear explosion; and we have seen that the number of expected unrelated pairs for $r<40\,\mu$ is only 0·2.

Fig. 13.



Distribution in the observed values of r for related pairs. The curves (A) and (B) show the expected distributions due to probable errors of measurement in the determination of the line of motion of the γ-ray of 0·25° and 0·5° respectively, the lifetime of the neutral mesons being assumed indefinitely short.

The number of events observed is not great enough, however, to allow us to put any weight on the indications of a finite lifetime given by the experiments, and we only note that if the apparent features can be confirmed by further observations of greater statistical weight, it may be possible to establish the magnitude of the lifetime if it is equal to or greater than 2×10^{-14} sec. The conclusion that the lifetime is less than 5×10^{-14} sec. appears to be established. It will be remarkable if the lifetime of the neutral particles is found to be just sufficiently great to be determined by this method, the only one available with the resources at present at our disposal.

10. The Intensity of the Neutral Mesons.

In order to determine the frequency of occurrence of the neutral mesons in comparison with the charged mesons, the total length, L, of the tracks of the shower particles in the scrutinized area round each star has been determined. Recent experiments indicate that about 80 per cent of these shower particles are π-mesons, and the contribution of fast protons can thus be excluded. Suppose that, on the average, the emission of each charged π-particle is accompanied by ϕ neutral mesons :

$$\mathcal{N}(\pi^0)/\mathcal{N}(\pi^{\pm})=\phi.$$

The length of path in the emulsion of the γ-rays formed by the decay of the neutral mesons is then given, without serious error, by

$$\frac{4}{5}\mathrm{L}\,.\,2\,.\,\phi=1{\cdot}6\mathrm{L}\phi.$$

The observation show that in this length of path, 15 pairs of electrons were created. Taking into account the small change with quantum energy of the " conversion-length ", the average length of path of a γ-ray in the emulsion before it produces a pair was calculated to be 4·6 cm., for the particular spectral distribution with which we are concerned (Heitler 1944). We may therefore write $\frac{1{\cdot}6\mathrm{L}\phi}{15}=4{\cdot}6$; with L=95 cm. it follows that

$$\phi=0{\cdot}45\pm0{\cdot}10.$$

The above result suggests that in the production of showers of charged particles, for every two charged mesons, one neutral particle is created. This conclusion receives support from a recent study by Camerini, Fowler, Lock and Muirhead (1950 b) in this laboratory, of the energy balance in the nuclear processes leading to the production of showers. These authors have compared the energy of the primary particles with that of the mesons and nucleons ejected in the disintegration. They have thus been able to show that there is an apparent disappearance of energy which can be accounted for if neutral π-particles are created. A similar result has been obtained by Green and Fretter (1950).

11. Conclusions.

We regard our experiments as proving that in the explosive disintegration of nuclei, neutral mesons are created in numbers equal to about half those of the charged π-particles. The mass of these neutral mesons is 295 ± 20 m_e, and they disintegrate into two γ-rays with a lifetime less than 5×10^{-14} sec. ; they may therefore be identified with the neutral π-mesons discovered in Berkeley. It is well known from the work of Steinberger (1949), and of Fukuda and Miyamoto (1949), that the mode of decay into two γ-rays indicates that the neutral meson is of spin zero.

The distribution in energy of the neutral mesons is closely similar to that of the charged mesons, and they arise in nuclear processes of the same type. They may therefore be regarded as having properties similar to those of neutral " heavy quanta " of Yukawa's theory.

The observed frequency of occurrence of the neutral mesons allows a simple explanation to be given of an observation by Rossi (1948) that the flows of energy represented by the " hard " and " soft " components of the cosmic radiation are of closely similar magnitude; an observation which appeared as an empirical fact at the time but which, as Rossi suggested, might have an important, deeper significance. Our results confirm the suggestion that both the " hard " and " soft " components arise in the nuclear disintegrations produced by protons and heavier particles of great energy, which are accompanied by the creation of charged and neutral π-particles. It appears to be unnecessary to attribute any of the soft component to " primary " electrons.

The lifetime of the π°-particles is certainly less than 5×10^{-14} sec.; and it will be difficult to determine it directly, and to establish the independent existence of the particles, with the methods at our disposal. A determination of the lifetime may, however, be possible by the present methods if it lies in the interval from 2×10^{-14} sec. to 5×10^{-14} sec.

Acknowledgments.

We wish to take this opportunity of thanking Professor C. F. Powell, F.R.S. for his continuous interest and encouragement in the course of this work. We are indebted to Dr. H. K. Heitler who prepared the diagrams, to Mr. R. A. Gattiker and Miss P. M. Dyer who carried out the photography and to Messrs. W. O. Lock and H. Muirhead for processing the emulsions. Without the work of Mrs. D. M. Ford, Mrs. J. Cowie, Miss J. Jones and Miss. M. Jones, who performed the arduous scrutiny of the plates with oil-immersion objectives, this investigation would not have been possible.

The authors were supported by grants from the Swedish Atomic Committee, the Department of Scientific and Industrial Research, and the Medical Research Council respectively.

APPENDIX.

For completeness, we outline the main features of the calculations. We have seen, that the energy of a quantum emitted at an angle θ with respect to the line of motion of a neutral meson of velocity $v=\beta c$, total energy $Bm_0c^2=\epsilon$ is given by the equation

$$h\nu=\mathrm{E}=\frac{m_0c^2}{2\mathrm{B}(1-\beta\cos\theta)}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

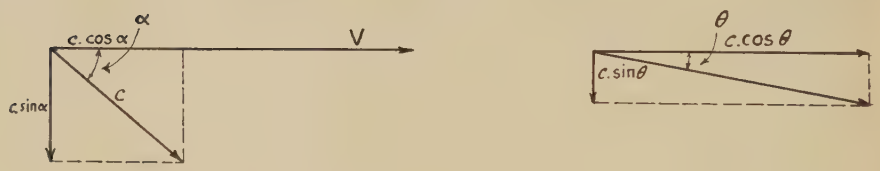
(a) *Angular Distribution of the Radiation.*

The quanta produced by the spontaneous decay of a homogeneous beam of neutral mesons may be assumed to be emitted isotropically in a coordinate system—the C-system—moving with the same velocity as the particles. The motions in this system are then transformed to the laboratory system—the L-system.

Suppose the angle of ejection of a quantum in the C-system, relative to the direction of motion of the meson, is α. The components of the velocity, c, of the quantum, parallel and perpendicular to the direction of motion of the particle, are $c \cdot \cos\alpha$ and $c \cdot \sin\alpha$, respectively (fig. 14). Let u_x and u_y be the corresponding components in the L-system. By the Einstein theorem for the addition of velocities we obtain

$$u_x = \frac{c\cos\alpha + v}{\left(1 - \frac{v\cos\alpha}{c}\right)}; \qquad u_y = \frac{c\sin\alpha}{B\left(1 - \frac{v\cos\alpha}{c}\right)},$$

Fig. 14.



where $B = 1/\sqrt{(1-\beta^2)}$. It follows that

$$\tan\theta = \frac{u_x}{u_y} = \frac{\sin\alpha}{B(\cos\alpha+\beta)}.$$

Writing $\sin\xi = \sqrt{(1-\beta^2)}$ and $\cos\xi = \beta$, it follows that :

$$\tan\theta = \frac{\sin\alpha\sin\xi}{\cos\alpha+\cos\xi}. \quad \ldots\ldots\ldots \quad (2)$$

From the identity $\sec^2\theta = \tan^2\theta + 1$, we then obtain the useful relations

$$\cos\theta = \frac{\cos\xi+\cos\alpha}{1+\cos\xi\cos\alpha}; \quad \ldots\ldots\ldots \quad (3)$$

and

$$\sin\theta = \frac{\sin\xi\sin\alpha}{1+\cos\xi\cos\alpha}; \quad \ldots\ldots\ldots \quad (4)$$

and from (2), (3) and (4)

$$\frac{d\alpha}{d\theta} = \frac{(\cos\xi+\cos\alpha)^2 \cdot \sec^2\theta}{\sin\xi(1+\cos\xi\cdot\cos\alpha)}. \quad \ldots\ldots\ldots \quad (5)$$

Let the intensity of the radiation in the L-system, measured as the number of quanta per unit solid angle, be $I(\theta)$; and in the C-system, $I(\alpha)$. We can then write

$$I(\theta) \cdot \sin\theta \cdot d\theta = I(\alpha) \cdot \sin\alpha \cdot d\alpha = K \cdot \sin\alpha \cdot d\alpha.$$

Putting $K=1$ for convenience, we obtain

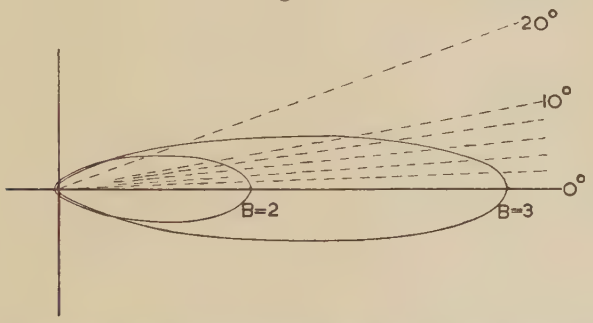
$$I(\theta) = \frac{\sin\alpha}{\sin\theta} \cdot \frac{d\alpha}{d\theta};$$

which reduces to

$$I(\theta) = 1/B^2(1-\beta\cos\theta)^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

The form of the distribution, for two values of B, is shown as a polar diagram in fig. 15.

Fig. 15.



Polar diagram showing the distribution in angle of the intensity of the γ-rays produced by the decay of homogeneous beams of neutral mesons of different energies.

(b) *Distribution in energy of the gamma-rays.*

We may write $N(E) \cdot dE = \frac{1}{2} I(\theta) \cdot \sin\theta \cdot d\theta$. From (1) and (6) we obtain

$$\frac{d\theta}{dE} = -\frac{2(1-\beta\cos\theta)^2}{\beta\epsilon(1-\beta^2)\sin\theta},$$

and, thence,

$$N(E) = 1/\beta\epsilon = \text{constant.} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

The distribution extends from

$$E_1 = (1-\beta)\epsilon/2 \qquad \text{to} \qquad E_2 = (1+\beta)\epsilon/2. \quad . \quad . \quad . \quad . \quad (8)$$

(c) *Properties of the γ-ray spectrum.*

We now consider an inhomogeneous beam of neutral mesons distributed in energy, the number in the interval from ϵ to $\epsilon+d\epsilon$ being $F(\epsilon) \cdot d\epsilon$. The γ-ray intensity due to the neutral mesons in this energy interval is given, according to (7) by the equation

$$\frac{F(\epsilon) \cdot d\epsilon}{\beta\epsilon} = \frac{F(\epsilon) \cdot d\epsilon}{\sqrt{(\epsilon^2-\epsilon_0^2)}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

between the limits

$$E_1=\tfrac{1}{2}\{\epsilon-\sqrt{(\epsilon^2-\epsilon_0^2)}\} \text{ and } E_2=\tfrac{1}{2}\{\epsilon+\sqrt{(\epsilon^2-\epsilon_0^2)}\}, \quad . \quad . \quad (10)$$

where ϵ_0 is the rest-energy of the neutral meson. At a γ-ray energy, E, the intensity is obtained by integrating equation (9) between the limits $\epsilon=E+\epsilon_0^2/4E$ and infinity ; for every neutral meson with energy above ϵ contributes to the γ-ray intensity at the energy E. This holds both for energies $E_1\leqslant\epsilon_0/2$ and $E_2\geqslant\epsilon_0/2$, because, when solving (10) for ϵ, we obtain in both cases

$$\epsilon=E_{1,2}+\epsilon_0^2/4E_{1,2}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

Thus the γ-ray spectrum is given by

$$N(E)=\int_{\epsilon}^{\infty}\frac{F(\epsilon)\,d\epsilon}{\sqrt{(\epsilon^2-\epsilon_0^2)}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

By differentiating (12) we obtain

$$F(E+\epsilon_0^2/4E)=E\left|\frac{dN}{dE}\right| \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

a relation which allows the distribution in energy of the mesons to be derived from the observed spectrum of the γ-radiation. At the quantum energy $E=\epsilon_0/2$, the intensity of the γ-radiation has a maximum value. It is clear from (13) that the slope, dN/dE, at this point is zero only when $F(\epsilon_0)=0$; *i. e.*, when there are no mesons of zero velocity in the beam.

(d) *The Rest Energy of the Neutral Meson.*

It follows from (12), which holds both for $E_1\leqslant\epsilon_0/2$ and $E_2\geqslant\epsilon_0/2$, that the intensity of the γ-radiation is the same for E_1 and E_2 when according to (11)

$$E_1+\epsilon_0^2/4E_1=E_2+\epsilon_0^2/4E_2$$

which gives

$$\epsilon_0^2=4E_1E_2$$

or

$$\epsilon_0=2\sqrt{(E_1E_2)}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (14)$$

This relation, which permits the derivation of the rest energy, ϵ_0, from the observed γ-ray spectrum, holds generally, being a result of the Doppler effect. It is therefore independent of the form of the " spectrum " of the neutral mesons.

(e) *Derivation of the r Distributions.*

Consider a meson emerging from a star (fig. 8), travelling a distance l before decaying. Assume that it then emits one of its γ-rays in the direction making an angle θ with the direction of motion of the neutral meson, so that the orthogonal distance from the star to the line of motion of the γ-ray is r. It follows that

$$r=l\,.\,\sin\theta\,;$$

and, as we have $N(\theta) \cdot d\theta = I(\theta) \cdot \sin\theta \cdot d\theta$, we may write

$$N(r) = N(\theta)\frac{d\theta}{dr} = I(\theta) \cdot \tan\theta/l.$$

Now $l = c\beta Bt$, where t is the time interval between the instants of creation and the decay of the neutral meson in the C-system. We assume that the meson decays according to the usual laws of radioactive decay, and then find that the number decaying between times t and $t+dt$ is

$$N(t)\,dt = \frac{N_0}{\tau_0}\exp(-t/\tau_0)\,dt,$$

τ_0 being the proper mean lifetime of the meson. Hence,

$$N(l)\,dl = \frac{N_0}{c\beta B\tau_0} \cdot \exp(-l/c\beta B\tau_0) \cdot dl\,;$$

and since

$$l = r/\sin\theta,$$

$$N(l)\,dl = \frac{N_0}{c\beta B\tau_0} \cdot \frac{r}{\tan\theta \sin\theta}\exp(-r/c\beta B\tau_0 \sin\theta)\,d\theta\,;$$

and

$$N(r) = N_0/c\beta B\tau_0 \cdot \int_0^{\pi} I(\theta)\exp(-r/c\beta B\tau_0 \sin\theta)\,d\theta.$$

We have performed this integration graphically for several values of the energy, and found that the calculated distributions in the values of r are nearly exponential for values greater than $10\,\mu$. For values $B \geqslant 2$ the curves are very nearly identical.

The final distribution is obtained by adding the curves for different values of B, each appropriately weighted according to the energy spectrum of the neutral mesons.

(f) *Background Pairs.*

The probability of finding an unrelated pair of which the bisector passes the star at a distance lying between r and $r+dr$ may be calculated from the known angular distribution and frequency of occurrence of such pairs in the plates.

Our observations show that γ-rays are isotropically distributed in the lower hemisphere (fig. 2). As a result of the geometry of the scanned area (fig. 10), two-thirds of the unrelated pairs point towards the star.

If $N(r)\,dr$ is the distribution in r formed by these pairs, we may write

$$N(r) = \tfrac{1}{4}\int_r^R I(\theta) \cdot \sin\theta \cdot \frac{d\theta}{dr} \cdot N(l) \cdot dl,$$

where l is the distance between the star and the origin of the pair, θ is the angle between the bisector and the line joining the origin of the pair with the star, and $N(l)\,dl$ is the number of pairs of which the origin lies

in the spherical shell of radii l and $l+dl$. Thus $N(l)\,dl=\rho 4\pi l^2\,dl$ where ρ is the observed number of pairs per unit volume. The angular distribution $I(\theta)$ is constant. R is the effective radius of the scanned area. Using the relation $r=l\,.\,\sin\theta$ we obtain

$$N(r)=\text{constant}\times\int_r^R l\,.\,\tan\theta\,d\theta$$

$$=\text{constant}\times r\sqrt{(R^2-r^2)}.$$

For $r\leqslant 600\,\mu$, we have $N(r)\sim\text{const}\,.\,R\,.\,r$, a relation which is true to within 10 per cent. The constant is then deduced from the observed density of " background " pairs and the known value of R.

References.

Bethe, 1934, *Proc. Camb. Phil. Soc.*, **30,** 524.

Bjorklund, Crandall, Moyer and York, 1950, *Phys. Rev.*, **77,** 213.

Bradt, Kaplon, and Peters, 1950, *Helv. Phys. Acta.*, **23,** 24.

Camerini, Coor, Davies, Fowler, Lock, Murihead and Tobin, 1949, *Phil. Mag.*, **40,** 1073.

Camerini, Fowler, Lock and Muirhead, 1950 a, *Phil. Mag.*, **41,** 413 ; 1950 b, *Ibid.* (in course of publication).

Cocconi, Loverdo and Tongiorgi, 1946, *Phys. Rev.*, **70,** 852.

Cocconi, 1949, *Rev. Mod. Phys.*, **21,** 26.

Chao, 1949, *Phys. Rev.*, **75,** 581.

Davies, Franzinetti and Perkins, 1950, *Phil. Mag.* (in course of publication).

Davies, Lock and Muirhead, 1949, *Phil. Mag.*, **40,** 1250.

Fowler, 1950, *Phil. Mag.*, **41,** 169.

Fretter, 1949, *Phys. Rev.*, **76,** 511.

Fukuda and Miyamoto, 1949 a, *Prog. Theor. Phys.*, **4,** III, 347 ; 1949 b, *Ibid.*, **4,** III, 391, 392.

Green and Fretter, 1950 (Private communication).

Gregory, Rossi and Tinlot, 1950, *Phys. Rev.*, **77,** 299.

Heitler, 1944, *Quantum Theory of Radiation*, 2nd Edition (Oxford : University Press).

King, 1950, *Nature, Lond.*, **165,** 526.

Modesitt and Koch, 1950, *Phys. Rev.*, **77,** 175.

Panofsky, Aamodt and York, 1950 (Private communication).

Piccioni, 1950, *Phys. Rev.*, **77,** 1.

Rossi, 1948, *Rev. Mod. Phys.*, **20,** 537 ; 1949, *Ibid.*, **21,** 104.

Steinberger, 1949, *Phys. Rev.*, **76,** 1180.

A large shower of mesons ($6 + 4p$) showing the origin of a related pair very close to the star. The pair has been marked by an arrow in the microphotograph.

Observer: Miss E. JAMES.

To face page 724

LXIV. CORRESPONDENCE

Gap Measurement as a Method of Analysing Cosmic Ray Stars in Photographic Emulsions.

By P. E. Hodgson*.

[Received May 31, 1950.]

Owing to the recent increase in sensitivity of photographic emulsions for nuclear research, tracks of particles of moderate energy are clogged, that is, the grains coalesce to form continuous blocks. Since grain counting is a well-known method of identifying and estimating the energy of such particles, it is desirable to extend the method to such tracks.

While in regions of moderate clogging some simple convention, such as one micron of solid track is equivalent to three grains, is satisfactory, this is highly artificial where the clogging is severe. In such regions it is natural to measure the total gap length per unit length, which is defined as the gap density. As grain counting is easier and more rapid than gap measurement, it is desirable to use it to the highest possible grain densities, which in practice are found to be in the region of 1·5 grains per micron, which corresponds to a gap density of about 0·1. Beyond this region gap measurement must be used.

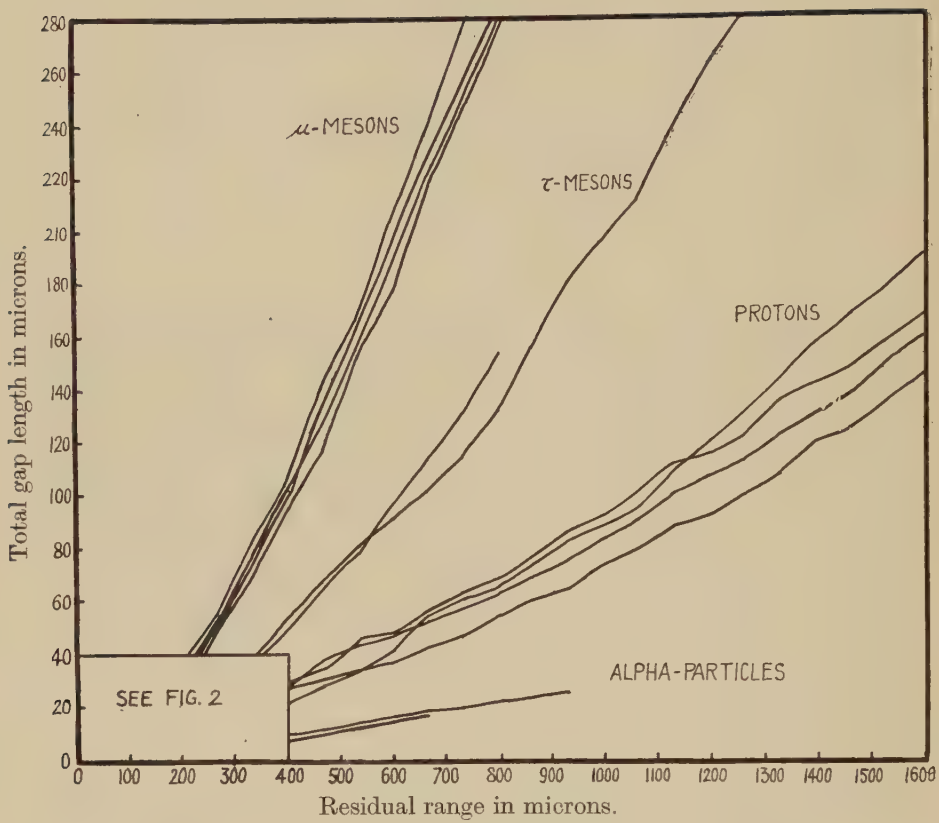
In order to relate the gap density to the energy loss, gap measurements versus residual range were made on the tracks of a number of particles of different mass and charge ending in the emulsion of an Ilford G5 plate exposed on the Jungfraujoch. Figs. 1 and 2 show the discrimination obtained by means of graphs of total gap length against residual range. If the track is longer than 400 microns, gap measurement gives almost certain identification, except between protons, deuterons and tritons, and between π- and μ-mesons. The two τ-mesons are those recently reported (Harding 1950).

The energy losses at points along the tracks were calculated from the known range–energy curves (Bradner *et al.* 1950), and hence a graph of gap density versus energy loss was constructed. These results are plotted on fig. 3, together with the standard deviations of measurements along 130 micron samples of track. In practice more than 130 microns of track are often available, and so the accuracy is higher than indicated by the figure. It is found that, above energy losses of 9 KeV./μ, the gap density and energy loss satisfy the empirical formula

$$\frac{dG}{dR}=0{\cdot}53\left(\frac{dE}{dR}\right)^{-0{\cdot}7},$$

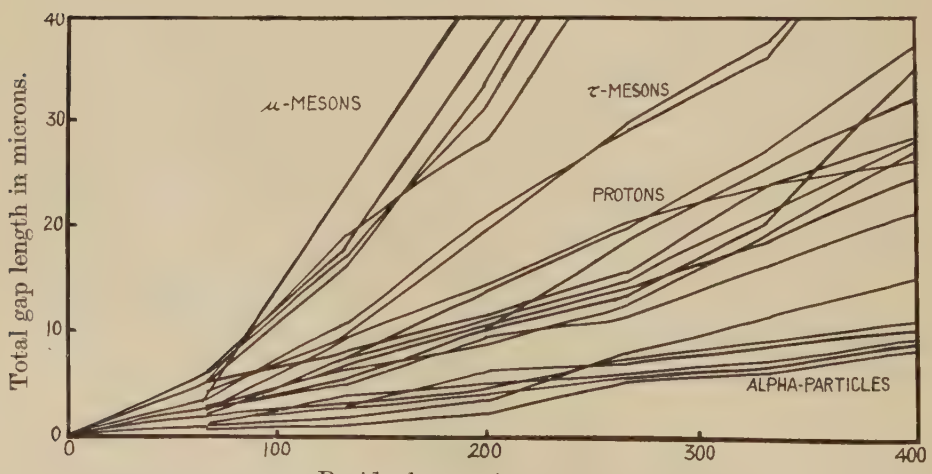
* Communicated by Sir George Thomson, F.R.S.

Fig. 1.



Total gap length against residual range for various particles.

Fig. 2.



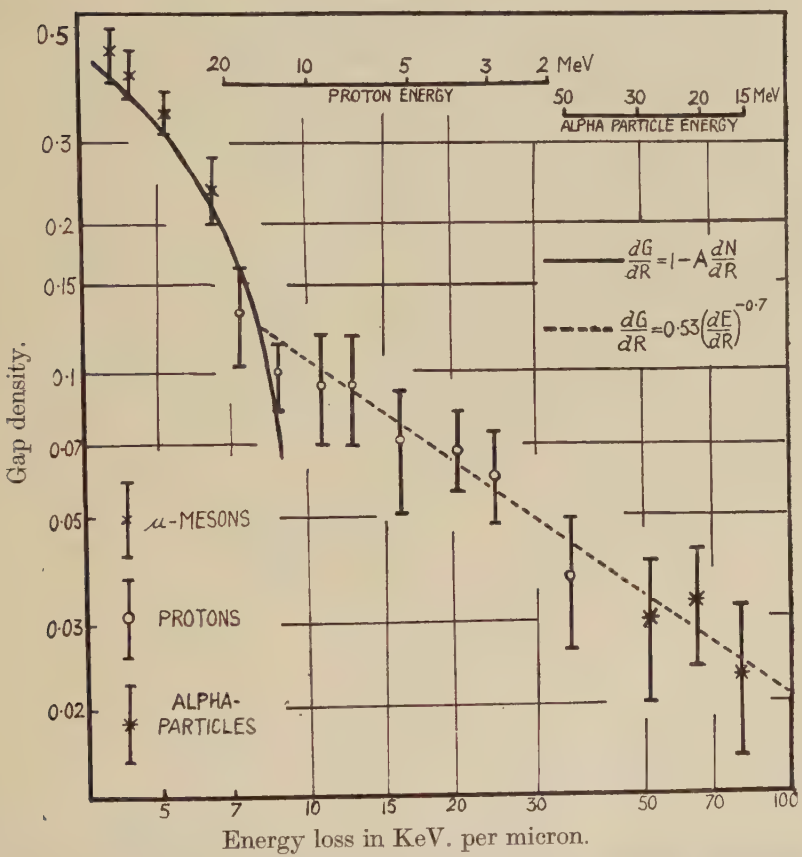
Total gap length against residual range for various particles.

where dE/dR is the energy loss in KeV./μ and dG/dR the gap density. The exact form of this expression depends on the type of emulsion and development, and so each batch of plates must be calibrated.

At lower grain densities, where the clogging is small, the gap density is given by

$$\frac{dG}{dR} = 1 - A\frac{dN}{dR},$$

Fig. 3.



Relation between gap density and energy loss with standard deviations of measurements along 130μ samples of track.

where dN/dR is the grain density in grains per micron empirically related to dE/dR, and $A \sim 0{\cdot}6\mu$ is the mean grain diameter. This curve is included in fig. 3 and agrees with the experimental points for high gap densities. It should be noted that the gap density is a pure number, and so, unlike the grain density, is independent of the size of the eye-piece scale divisions.

The procedure to be followed in order to identify and measure the energy of those particles from cosmic ray stars in electron sensitive emulsions producing clogged tracks depends on whether the particle under investigation ends in the emulsion or not.

If the track ends in the emulsion, a gap measurement is made and the particle identified by comparison with figs. 1 and 2. Visual examination of the track for scattering and δ-rays (Sörensen 1949) confirms the identification. Use of range–energy curves gives the energy of emission of the particles from the star.

If the track does not end in the emulsion, the gap density identifies the particle, provided the track is sufficiently long, except perhaps for some uncertainty between a slow meson and a fast proton, and between a slow proton and a fast alpha-particle. Inspection of scattering in the first case and δ-rays in the second, resolves this difficulty. Using fig. 3, the energy loss can be found from the gap density and hence the energy is known from the energy–energy loss curves.

The above analysis can only be carried out if the track makes a small angle with the plane of the emulsion. The correction factor which must be applied to the gap densities of dipping tracks may be derived from a simple geometrical model. It may be shown that a gap of length g appears of length $(g+\mathrm{A})\cos\theta'-\mathrm{A}$ when viewed from a direction making an angle θ' with the normal to the track. Assuming that the distribution of gap lengths may be replaced by a mean gap length g_m and that the lengths of blocks of grains remains constant during shrinkage, it may further be shown that

$$\frac{d\mathrm{G}}{d\mathrm{R}}=\frac{\cos\theta}{\cos\theta'}\left(1+\frac{\mathrm{A}(1-\cos\theta')}{g'_m}\right)\left(\frac{d\mathrm{G}}{d\mathrm{R}}\right)'+1-\frac{\cos\theta}{\cos\theta'},$$

where $\tan\theta/\tan\theta'$=shrinkage factor$\simeq 2{\cdot}7$,

$\left(\frac{d\mathrm{G}}{d\mathrm{R}}\right)'$=apparent gap density,

g'_m=mean apparent gap length.

For $\theta<10^\circ$ this correction is small, while for $\theta>20^\circ$, gap measurement is unreliable. The correction factor is large for tracks of low gap density, since the mean gap length decreases with gap density.

This method, although more laborious, is better than simply counting the number of gaps because it gives better discrimination between particles and also it is less sensitive to the angle of dip of the track.

References.

Bradner, H., Smith, F. M., Barkas, W. H., and Bishop, A. S., 1950, *Phys. Rev.*, **77**, 462.

Harding, J. B., 1950, *Phil. Mag.*, **41**, 405.

Sörensen, S. O. C., 1949, *Phil. Mag.*, **40**, 947.

A Note on the Use of Resistance Thermometers for Measurement of Rapidly Changing Temperatures.

By C. B. Daish, D. H. Fender and A. J. Woodall.
Military College of Science, Physics Branch *.

* Communicated by the Authors.

[Received June 2, 1950.]

In a recent issue of the *Philosophical Magazine* (1950, **41,** 468–477), Leah, Rounthwaite and Bradley describe the use of resistance thermometers in studying the high temperatures produced in explosions. These investigators sought to follow a temperature which rises more than 1000° C. in 20 or 30 milliseconds by using a platinum alloy wire of length 1 inch and diameter 0·0005 inch, and to eliminate radiation errors by a calibration experiment designed to disentangle natural convection from radiation.

In investigations we are making, we have followed the temperature-changes of a long, fine platinum wire suddenly heated and then allowed to cool in an enclosure under conditions in which conduction through the gas is the dominant method of transference of heat. In such conditions the cooling curve of the wire is roughly exponential and experiment substantiates our theory, which assigns to the cooling process a time-constant depending on the thermal conductivity of the gas, on the thermal capacity per unit length of the wire and on the geometry of the system, including the dimensions of the containing vessel. The plain wire used by these investigators would display in such conditions a time-constant of the order of 20 millisec., and the quartz-coated wire would behave appreciably differently. The relatively stout leads would have a time-constant of the order of half a minute, but this will be considerably reduced if natural convection is allowed to come into play.

If this system of wires and leads is exposed to a sudden change in temperature, since convection currents take time to form, the initial heat transfer will be solely by radiation and conduction through the gas and along wires. Conduction to the leads is not negligible. Indeed a calculation shows that in the extreme case when the surface of the wire loses heat by conduction alone, the average temperature of the wire only rises by 93 per cent of the rise at the centre, followed by a subsequent slow drift as the leads warm up. Since the end effects will be very different in the main and the calibration experiments, an error will be thereby introduced, and the effect will moreover be appreciably different for the quartz-coated wire.

Further, the attack on the radiation correction in the explosion experiments relies on a subsidiary experiment on natural convection. As mentioned above, natural convection can have little bearing on the critical first stage of the explosion experiments and the method of obtaining a radiation correction is therefore invalid.

Lastly, but probably of most importance, the true shape of the temperature–time curve must be badly distorted by the lag in response of the wire. The fact that the plain and the quartz-covered wires have different time-constants will ensure that the distortions are not the same in the two cases and this factor alone might conceivably account for the differences observed in the experiments. Certainly the effect of this factor must be very carefully considered before any reliance can be placed on deductions from the curves presented.

It was the conviction that the response of resistance thermometers and thermocouples to sudden changes warranted more attention which prompted us to undertake our investigations some time ago, and we hope shortly to be in a position to make a more adequate communication on the subject.

May 22nd, 1950.

LXV. *Notices of New Books and Periodicals received.*

Relations entre les phénomenes solaires et géophysiques. Méthodes de calcul dans des problèmes de mécanique. Cinétique et mécanisme des réactions d'inflammation et de combustion en phase gazeuse. (Centre National de la Recherche Scientifique : Paris.)

THESE are records of three of the international colloquia which have been organized in recent years with the help of the French Scientific Research Centre.

The first, held at Lyons in 1947, dealt with solar and geophysical phenomena. A main problem in this field is to establish the complex relations between the incident solar radiations and effects observable at the earth's surface. Some evidence is obtainable by balloon and, more recently, by rocket flights, but most of the experimental material presented here is of radio observation on the ionosphere and optical investigation of such phenomena as the aurora and the luminosity of the night sky. On the theoretical side, there is a paper by S. K. Mitra on unsolved problems in the field and articles by H. S. W. Massey and Marcel Nicolet on the photochemistry of the high atmosphere.

The second volume is divided into two parts referring to meetings at Marseilles and Paris in April 1948. The first part contains an exposition by D. N. de G. Allen of Southwell's relaxation method, and an account by T. Vogel of the escalator method of Morris and Head for the determination of normal vibrations. The second is devoted to problems of fluid flow, including in particular articles by J. M. Burgers on turbulent flow, and by L. Malavard on the use of electrical analogues.

The third volume records the colloquium on the initiation and propagation of combustion in gases which was held in Paris at the end of April 1948. It is concerned for the most part with the kinetics and detailed molecular mechanism of the reactions involved. It would give a wrong impression to cite individual articles. Considerable advance has been made in this subject during and since the war, and this colloquium contains much interesting and valuable information.

The benefit of colloquia such as these is greatest for those fortunate enough to take part, but these printed records will be of interest to many not so experienced in such special field as to have contributed thereto. G. W.

The Chemical Elements and their Compounds. By N. V. Sidgwick. Two vols. [Pp. xxxii+1703.] (Oxford: Clarendon Press, 1950.) Price 70*s*.

This work stresses the basic principles underlying the reactions and properties of the chemical elements and compounds. A simple and sensible layout based on the classification of the periodic table, and the emphasis placed on the relations between elements of a group and the groups themselves, provide a clear and satisfying picture of chemistry as a coherent whole. Perhaps the most striking feature of the discourse is the rational treatment accorded to the organic and metal-organic compounds, which are carefully fitted into the pattern instead of being relegated to more highly specialized works. Technical details are reduced to a minimum, with the result that the fundamental structure of chemistry is thrown into relief. Clear exposition combines with an extensive bibliography to make these two volumes a necessity for the chemist's library.

D. H. M.

Non-Linear Vibrations in Mechanical and Electrical Systems. By J. J. Stoker. [Pp. xix+273.] (Interscience Publishers, Inc.) Price 40*s*.

The author is concerned with the free and forced oscillations in two ordinary non-linear differential equations, and although there is no claim to give a survey of the work of the past twenty years, he does take account of much recent work and includes an up-to-date bibliography.

After a discussion of the stability of the first order homogeneous differential equation near a singularity, he considers the forced oscillation in Duffing's Equation, in which a small non-linear element occurs in the restoring force. Sections are included on jump phenomena, that is, discontinuities in the solution as the period of the forcing term is varied and on the appearance of subharmonic oscillations of a permanent type.

The other differential equation is Van der Pol's, in which the damping term is non-linear, so that the damping is negative for small velocities and positive for large velocities. Self-excited oscillations are considered first of all and then the stability of the periodic solutions arising from a small forcing term. A chapter is included on Mathieu's equation and its application to the stability of the solution of Duffing's equation.

The presentation up to this point is elementary and within the reach of non-mathematicians, but to conclude the book there are six appendices proving some results of a more abstruse nature.

K. S.

Les spectres de rayons X et la structure électronique de la matière. By Yvtete Cauchois. (Paris: Centre National de la Recherche Scientifique.)

This short book contains an admirable survey of a subject to which the school of Mlle. Cauchois in Paris has made important contributions. The first part gives a rapid conspectus of the field and discusses especially the influence of chemical binding on the position of the discontinuities in X-ray absorption spectra, while the remaining and greater part of the book deals with the interpretation of the fine structure of these spectra in terms of the permitted electronic energy levels, especially for crystalline absorbers. Emission spectra are also fully treated, though in this field there is a much smaller mass of reliable experimental work.

This subject is of the highest importance as giving substantial information on the behaviour of electrons in molecules and condensed matter, and it is an especial merit of the book to draw attention to the decidedly unsatisfactory state of theoretical work on these questions. Many of the unsolved problems raised are of considerable interest. This book should be read by all interested in the electronic structure of matter.

G. W.

Dynamics of real Fluids. By E. G. Richardson. [Pp. 143.] (London: Edward Arnold, 1950.) Price 21s.

There has for some time been a need for an up-to-date text-book on the physics of real fluids, avoiding both the mathematical complexity of Lamb's classical treatise on ideal fluids, and the engineering details of, say, Dodge and Thomson. One hoped that Dr. Richardson's book would meet this need, and a glance at the chapter headings seemed encouraging. A short introduction on classical hydrodynamics is followed by a long chapter on "Fluids of Small Viscosity" which deals largely with boundary layer theory and turbulent wakes. There follow chapters on compressible fluids, fluids with free surfaces, fluids with anomalous viscosities, and problems involving thermal effects.

As a text-book for students the book seems to the present reviewer to have serious faults. As a summary of the present state of knowledge on the topics treated, the book is probably unique: but to the non-specialist it is very difficult reading. The author's very familiarity with his subject-matter is probably a contributory cause—his researches figure in almost every chapter. The treatment attempted is largely non-mathematical, but equations are sufficiently numerous to necessitate an attempt to understand them, and insufficiently detailed to assist the understanding. The exposition seems too rather obscure in places, and the number of mistakes and misprints is such that more than twenty were noticed at a single reading.

In short the book is excellent in conception, but the execution leaves much to be desired. N. T.

An Introduction to Heat Transfer. By M. Fishenden and O. A. Saunders. [Pp. 205.] (Oxford: Clarendon Press, 1950.) Price 15*s*.

This book is intended primarily for engineering students but others will also find it a useful summary of much information of practical value. It attempts to explain the fundamental physical principles of the three mechanisms of heat transfer, conduction, convection and radiation, and at the same time to provide quantitative formulæ and numerical data in a form suitable for practical applications. In most cases the basic formulæ are quoted without proof, but are followed by an adequate comparison with available experimental data. Each chapter contains a number of worked examples to illustrate the use of the formulæ.

The first third of the book deals with radiation and conduction in two chapters of about equal length, whilst the remainder is devoted to problems of natural and forced convection. The chapter on thermal conduction contains an account of Schmidt's graphical method of solving problems of unsteady heat flow, and emphasizes the importance of numerical and analogue methods generally. The treatment of thermal convection is based on formulæ which satisfy the principle of dynamical similarity, and is preceded by an elementary discussion of the application of dimensional analysis to convection problems.

The present reviewer feels that it is a valid criticism that this book attempts to cover too large a field in a limited space. In consequence, the treatment of some important aspects of the heat transfer problem is rather brief; for example, forced convection to a gas flowing at high speed is covered in a single paragraph. E. P. Hicks.

ERRATA.

In the article by K. C. Westfold, of the June issue, there should be added to the References:

Smerd, S. F. and Westfold, K. C., 1949, *Phil. Mag.*, **40**, 831–48.

[*The Editors do not hold themselves responsible for the views expressed by their correspondents.*]